Stratosphere - Troposphere Interaction during Stratospheric Sudden Warming Events

by

Daniela I.V. Domeisen

Submitted to the Department of Earth, Atmosphere and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2012

Author .....
Department of Earth, Atmosphere and Planetary Sciences
September 19, 2012

Certified by .....
R. Alan Plumb
Professor of Meteorology
Thesis Supervisor

Accepted by .....
Robert D. van der Hilst
Schlumberger Professor of Earth Sciences
Head, Department of Earth, Atmospheric and Planetary Sciences

Stratosphere - Troposphere Interaction during Stratospheric Sudden Warming Events

by

Daniela I.V. Domeisen

Submitted to the Department of Earth, Atmosphere and Planetary Sciences
on September 19, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The stratosphere and the troposphere exhibit a strong coupling during the winter months. However, the coupling mechanisms between the respective vertical layers are not fully understood. An idealized spectral core dynamical model is utilized in the present study in order to clarify the coupling timing, location and mechanisms. Since the coupling between the winter stratosphere and troposphere is strongly intensified during times of strong stratospheric variability such as stratospheric warmings, these events are simulated in the described model for the study of stratosphere - troposphere coupling, while for comparison the coupling is also assessed for weaker stratospheric variability.

While the upward coupling by planetary-scale Rossby waves in the Northern Hemisphere is well understood, the Southern Hemisphere exhibits traveling wave patterns with a weaker impact on the stratospheric flow. However the tropospheric generation mechanism of these waves is not well understood and is investigated in this study. It is found that in the model atmosphere without a zonally asymmetric wave forcing, traveling waves are unable to induce a significant wave flux into the stratosphere. In the absence of synoptic eddy activity, however, the tropospheric flow is baroclinically unstable to planetary-scale waves, and the generated planetary waves are able to propagate into the stratosphere and induce sudden warmings comparable in frequency and strength to the Northern Hemisphere. While baroclinic instability of long waves may be further strengthened by the addition of moisture, the real atmosphere also exhibits strong synoptic eddy activity, and it will have to be further explored if the atmosphere exhibits periods where synoptic eddies are weak enough to allow for baroclinic instability of long waves.

In order to further investigate the coupling between the stratosphere and the troposphere, cases of strong coupling are investigated in the analysis of a Northern Hemisphere - like winter atmosphere. A realistic frequency and strength of sudden warmings is obtained using a zonal wave-2 topographic forcing. An angular momentum budget analysis yields that the Eliassen-Palm (EP) flux is closely balanced by the residual circulation dominated by the Coriolis term on a daily basis, while the

change in zonal wind is a small residual between these dominant terms. In the stratosphere, the EP flux term and the Coriolis term balance well in time but not exactly in magnitude, yielding a polar stratospheric weakening of the zonal mean wind as observed during stratospheric warmings. In the troposphere, the loss of angular momentum before a sudden warming induces a weak negative annular mode response, which is amplified by the downward propagating signal about three weeks after the sudden warming. The angular momentum budget does not reveal the mechanism of downward influence, but it nevertheless clarifies the momentum balance of the stratosphere - troposphere system, indicating that the effects of the waves and the residual circulation have to be considered at the same time.

Since the annular mode response cannot be directly investigated using the angular momentum budget, the annular mode coupling between the stratosphere and the troposphere is further investigated using a statistical approach. The annular mode response is often framed in terms of Empirical Orthogonal Functions (EOFs), but it is here found that for the stratosphere - troposphere system with its strong vertical pressure gradient, EOFs are strongly dependent on the weighting of the data, while Principal Oscillation Patterns (POPs) are considerably less sensitive to an applied weighting while returning the dominant structures of variability. This encourages further research and application of POP modes for the use of stratosphere - troposphere coupling.

These findings represent an improvement of the understanding of stratosphere - troposphere coupling and the results are another step in the direction of finding the mechanism of stratosphere - troposphere coupling and the downward influence after the occurrence of a stratospheric sudden warming, which may influence long-term weather prediction in the troposphere.

Thesis Supervisor: R. Alan Plumb
Title: Professor of Meteorology

Acknowledgments

I would like to first and foremost thank my parents for providing me with the financial freedom and moral support to pursue my studies in the United States. Also, I am greatly thankful to my boyfriend Jens Vogelgesang for his great support, his frequent visits and efforts to come to the US.

For my research work I have received great support from my advisor Alan Plumb. I would like to thank him for guiding me through the PhD and for taking whatever time was needed to help me with any research question I had. In addition I would like to thank my committee members: Kerry Emanuel for encouraging me to come to MIT, for serving as the thesis committee chair, and for helpful discussions, and Paul O'Gorman and Paul Kushner for helpful discussions throughout my PhD. I would also like to thank all the students and postdocs within my research group, especially Michael Ring, Cegeon Chan, Gang Chen and Eric Leibensperger for frequent helpful discussions of the research results as well as modeling questions. In addition, I appreciated the strong network among my advisor's former students and postdocs, especially Lorenzo Polvani and Volkmar Wirth. I would also like to thank the research groups at ETH Zürich for hosting me every summer throughout my PhD, especially Christoph Schär and Heini Wernli, as well as Rahel Buri for her great administrative efforts during my stays.

In addition I would like to thank Mary Elliff and Roberta Allard for truly helpful administrative work and for going out of their way in order to be able to help, even if not officially in charge. I would also like to thank Greg Shomo for being enormously helpful with computing and cluster questions, for being available and for always smiling despite his very busy work schedule. I would also like the MIT Weather Team for their support, enthusiasm and for everything I have learned being on the team, especially Lodovica Illari.

I would further like to thank the other students in the program, especially the students in my year, Diane Ivy, Malte Jansen, Anita Ganesan and Neil Zimmerman, as well as other students in the program, especially Martha Buckley for being a great

mentor, as well as Holly Dail, Ryan Abernathey, Allison Wing and many more for the great support and friendship during all these years at MIT. I would also like to thank my various office mates over the years, especially Yang Zhang for great advice and Stephen Messenger for many interesting and helpful discussions. I would also like to thank my friends and roommates for their valued support and assistance.

Outside of my research work, I would like to thank the former and current members of the SwissLinkBoston board, Yvonne Moret, Stefanie Frischknecht, Karl Külling, Aron Bläsi, and Valentin Splett for the great support and shared leadership during the build-up, maintenance and handover of the SwissLinkBoston network. In addition, I would like to thank the Swiss Consulate swissnex for the great collaboration and for their support, as well as the other Swiss organizations in the Boston area.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The stratosphere, the atmospheric layer between about 10 and 50 km height above the Earth's surface, has recently been attracting an increasing interest due to the finding that the stratosphere is able to actively interact with the troposphere, instead of the stratosphere merely responding to tropospheric forcing from below. Baldwin and Dunkerton (2001) have shown that events of large stratospheric variability are followed by tropospheric weather anomalies for up to two months after an event. Due to the comparatively longer memory of the lower stratosphere relative to the troposphere, signals of strong stratospheric variability may help increase subsequent tropospheric predictability [e.g. Orsolini et al. (2011)]. On longer timescales, Polvani et al. (2011) have shown that stratospheric ozone depletion has a significant impact on surface climate, even dominating the effect of climate change in the Southern Hemisphere. Understanding stratospheric variability is therefore crucial to better be able to predict tropospheric weather and climate.

The stratospheric circulation is to first approximation an interplay between radiative processes involving solar heating and chemical tracers, dominantly ozone, which are transported through the stratosphere, as well as dynamical processes involving wave propagation and wave breaking. The stratospheric *polar vortex*, a strong circumpolar flow in the winter stratosphere, can reach wind speeds of up to $90\,\mathrm{ms}^{-1}$

in Southern Hemisphere winter, with smaller average speeds on the order of $50\,\mathrm{ms}^{-1}$ in Northern Hemisphere winter. While radiative processes maintain the polar vortex, planetary-scale *Rossby waves*, which are dominantly caused by longitudinally asymmetric heating and topography at the Earth's surface, are responsible for strong disruptions of the vortex in so-called *Stratospheric Sudden Warming* events. These warmings are sudden (on time scales of days compared to radiative time scales of several weeks) changes in the stratospheric flow and temperature patterns. No definite answer has yet been given about their cause, although they were discovered and described more than 50 years ago [e.g. Scherhag (1952), Scherhag (1953), Labitzke (1981)].

The research performed for this thesis is mainly geared towards a better understanding of the conditions that lead to stratospheric sudden warmings as well as the connections to tropospheric variability under both quiet and sudden warming conditions. The co-variability and connection between the stratosphere and the troposphere is studied in terms of both the tropospheric origins and stratospheric consequences of wave propagation. It will be shown that it is not possible to consider wave propagation separately from the response of the atmosphere to these propagating and dissipating waves, as this response is observed on time scales shorter than the propagation of the waves. Both an angular momentum budget as well as a statistical approach are taken in order to better characterize the connections between the troposphere and the stratosphere. Sudden warming events are illuminating examples of anomalous vertical coupling between the stratosphere and the troposphere, which makes them ideal candidates for the study of strong coupling. While it is better understood why the coupling is weak during the summer season, it is illuminating to study winter seasons exhibiting weak coupling in comparison to winters with strong stratospheric variability. Both weak and strong coupling for winter conditions are examined in this thesis.

The remainder of this chapter will introduce the dynamics of the stratosphere and how it can be framed, and it will give an introduction to stratospheric warmings as well as more general stratosphere - troposphere coupling during the winter season.

1.2 Dynamics of the Stratosphere

In the stratosphere, many dominant processes act on comparatively slow time scales as compared to tropospheric processes: small Rossby waves are filtered out at lower levels to leave only planetary-scale waves to propagate into the upper stratosphere (Charney and Drazin, 1961), and in addition water vapor, a dominant factor in tropospheric weather and climate, has a much smaller influence on the stratospheric circulation. Despite its relative simplicity, the stratosphere exhibits various unique and interesting phenomena such as the quasi-biennial oscillation (connected to gravity waves) and stratospheric sudden warmings (connected to Rossby waves).

The stratosphere is dominantly heated from above by absorption of short-wave radiation by ozone. Ozone is most abundant with respect to air density in the upper stratosphere and slowly decreases in abundance with height. Ozone heating is the main factor responsible for the vertical temperature structure of the stratosphere, with temperature increasing with height throughout the stratosphere, while the troposphere is heated from below through long-wave radiation emitted from the surface which is itself heated by short-wave radiation from the sun. This vertical heating profile makes the stratosphere strongly stable in the vertical, explaining the slow time scales and the absence of convection.

One of the dominant stratospheric time scales is the seasonal cycle, since the mean state of the stratosphere is strongly dependent on available sunlight: during polar night no sunlight reaches the polar regions, leading to a strong meridional temperature gradient between the summer and winter hemisphere, which is particularly strong at the edge of the polar night. The polar vortex (which for this reason is also called the *polar night jet*) emerges from this strong gradient: it forms in fall and reaches its maximum strength in midwinter, but weakens to a complete wind reversal in late spring. Winds are weaker and easterly throughout the summer stratosphere. The midwinter vortex further isolates the polar night regions from the midlatitudes, thereby limiting the transport of e.g. ozone towards the pole. The stratosphere is thus dominated by a very strong seasonal cycle.

As will be shown in the next section, the presence of waves introduces variability on shorter time scales to the winter stratosphere: the time scale on which waves are able to propagate and interact with the mean flow (on the order of days) and the radiative time scales on which the flow recovers to radiative equilibrium (on the order of several weeks). This induces fundamental differences between the Southern and the Northern Hemisphere in terms of their respective winter variability [e.g. Plumb (1989)]: The Southern Hemisphere winter exhibits a much weaker variability of the polar vortex, which grows to significantly greater midwinter wind speeds. This behavior is generally attributed to the stronger stationary Rossby wave forcing at the surface in the Northern Hemisphere due to its strong topography and the zonally asymmetric stationary heating caused by the distribution of land and ocean in the extratropical Northern Hemisphere. The Southern Hemisphere exhibits a much smaller stationary forcing component, mostly due to Antarctic topography, while the Southern Ocean exhibits a dominantly zonally symmetric heating pattern. The Southern Hemisphere stratosphere is dominated by traveling waves which are generated in the troposphere and which in general are not able to induce strong stratospheric variability. Chapter 4 will investigate the generation of traveling waves for the Southern Hemisphere mean state.

The large wave amplitudes throughout Northern Hemisphere winter are able to consistently weaken the polar vortex, thereby keeping the vortex from reaching wind speeds comparable to its Southern Hemisphere counterpart. In the Southern Hemisphere, on the other hand, vortex wind speeds grow continuously into winter and wave amplitudes exhibit a minimum in midwinter (Hirota et al., 1983), indicating that wind speeds may be too strong even for zonal wave-1 to propagate (Plumb, 1989).

1.2.1 Formal Description of Stratospheric Variability

Extratropical stratospheric dynamics can be described by a set of primitive equations using the quasi-geostrophic assumptions of small Rossby number and low frequency motion, while neglecting friction, which is small in the stratosphere. The dynamics

of the zonally averaged state can then be formally separated from the dynamics of the deviations from the zonal mean state, while including a characterization of the interaction between them. This is achieved by separating all state variables in the set of primitive equations into zonal mean and wave disturbance parts, leading to a description of e.g. the flow in terms of the two-way interaction of the zonal mean flow with the waves (i.e. the deviations from the zonal mean flow). This separation leads to the Eulerian mean equations [e.g. Andrews et al. (1987), Equations 3.3.2a-e], a set of primitive equations describing the effect of an eddy forcing on the mean state of the atmosphere, and vice versa. However, these equations do not readily clarify the connection between the fluxes of heat and momentum and the zonal mean state. In order to be able to obtain a set of equations which combines the wave forcing into a single term in the momentum equation, a residual circulation can be defined according to

$$\overline{v}^* = \overline{v} - \frac{1}{\rho_0}\frac{\partial}{\partial z}\Big(\rho_0 \frac{\overline{v'\theta'}}{\overline{\theta}_z}\Big) \tag{1.1}$$

$$\overline{w}^* = \overline{w} + \frac{1}{a\cos\varphi}\frac{\partial}{\partial \varphi}\Big(\cos\varphi \frac{\overline{v'\theta'}}{\overline{\theta}_z}\Big), \tag{1.2}$$

where subscripts denote derivatives, i.e. $\overline{\theta}_z \equiv \frac{\partial \overline{\theta}}{\partial z}$. This definition leads to the Transformed Eulerian Mean (TEM) equations [e.g. Andrews et al. (1987), Equations (3.5.2a-e)] given in spherical log-pressure coordinates by

$$\frac{\partial \overline{u}}{\partial t} + \overline{v}^*\Big[\frac{1}{a\cos\varphi}\frac{\partial}{\partial\varphi}(\overline{u}\cos\varphi) - f\Big] + \overline{w}^*\frac{\partial \overline{u}}{\partial z} - \overline{X} = \frac{1}{\rho_0\, a\cos\varphi}\,\nabla\cdot F, \tag{1.3}$$

$$\overline{u}\Big[f + \frac{\overline{u}}{a}\tan\varphi\Big] + \frac{1}{a}\overline{\Phi}_\varphi = G, \tag{1.4}$$

$$\overline{\Phi}_z - \frac{R}{H}\,\overline{\theta}\, e^{-\kappa z/H} = 0, \tag{1.5}$$

$$\frac{1}{a\cos\varphi}\frac{\partial}{\partial\varphi}(\overline{v}^*\cos\varphi) + \frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0\overline{w}^*) = 0, \tag{1.6}$$

$$\overline{\theta}_t + \frac{\overline{v}^*}{a}\overline{\theta}_\varphi + \overline{w}^*\overline{\theta}_z - \overline{Q} = -\frac{1}{\rho_0}\frac{\partial}{\partial z}\Big[\rho_0\frac{\overline{v'\theta'}\,\overline{\theta}_\varphi}{a\overline{\theta}_z} + \overline{w'\theta'}\Big] \tag{1.7}$$

where $\rho_0 = \rho_0(z)$, Φ is geopotential, X and Q represent friction and diabatic heating, and G encompasses deviations from gradient wind balance. The vertical coordinate z is a log-pressure coordinate as defined in Section 5.3. The EP flux $\mathbf{F} = (0, F^\varphi, F^z)$ in the meridional plane is given by

$$F^\varphi = \rho_0 \, a \cos\varphi \Big[\frac{\partial \overline{u}}{\partial z} \frac{\overline{v'\theta'}}{\overline{\theta}_z} - \overline{v'u'} \Big] \tag{1.8}$$

$$F^z = \rho_0 \, a \cos\varphi \Big[\Big(f - \frac{1}{a\cos\varphi} \frac{\partial}{\partial \varphi} (\overline{u} \cos\varphi) \Big) \frac{\overline{v'\theta'}}{\overline{\theta}_z} - \overline{w'u'} \Big] \tag{1.9}$$

and the divergence of the EP flux is given by

$$\nabla \cdot F = \frac{1}{a\cos\varphi} \frac{\partial (F^\varphi \cos\varphi)}{\partial \varphi} + \frac{\partial F^z}{\partial z} \, . \tag{1.10}$$

All other variables are defined in Appendix A and in Andrews et al. (1987). Equation (1.3) is the momentum equation governing the temporal change in the zonal mean zonal wind, Equation (1.4) indicates geostrophic balance, Equation (1.5) is the hydrostatic equation, Equation (1.6) is the continuity equation and Equation (1.7) is the thermodynamic equation describing the response to a heating.

The effect of a momentum forcing on the atmosphere can now be described more readily: The eddy forcing terms now do not appear in the thermodynamic equation anymore, yielding a more straight forward characterization of the influence of a wave forcing onto the atmosphere, and illustrating that the eddy momentum flux and the eddy heat flux act together to yield a change in the mean flow. The TEM momentum equation (Equation (1.3)) illustrates how a wave forcing, given by a convergence of EP flux $-\nabla \cdot F$, may induce both a deceleration of the zonal mean zonal wind $\overline{u}$ as well as a residual circulation in the latitude/height plane given by $(\overline{v}^*, \overline{w}^*)$. The temporal change in zonal mean wind $\overline{u}$ is directly related to a temporal change in relative angular momentum per unit mass, given by $\overline{u} \, a \, cos\varphi$.

An additional benefit of framing stratospheric dynamics using the TEM equations is the emergence of a conservation law for wave propagation. Eliassen and Palm (1961) found that for steady, linear waves on a zonal flow, if friction and heating

can be neglected, that the divergence of the EP flux is zero. This leads to the wave activity density A as a conserved quantity in

$$\frac{\partial A}{\partial t} + \nabla \cdot F = D,$$

[e.g. Andrews et al. (1987), Equation (3.6.2)] where D represents diabatic and frictional effects. This conservation law illustrates the need for transient or frictional effects in order for a non-zero divergence of the EP flux to occur.

The description of stratospheric dynamics by the TEM equations will help frame and clarify the effect of a wave forcing onto the extratropical stratosphere. Such wave forcings are especially strong during strong upward wave propagation from the troposphere and during stratospheric sudden warmings in particular, as introduced in the next section.

1.3 Stratospheric Sudden Warmings

The propagation of Rossby waves, which have been identified as the major driver of stratospheric sudden warmings, is governed by a propagation law (Charney and Drazin, 1961) which inhibits wave propagation if the phase speed of the waves exceeds the background wind speed. The Rossby waves which are dominantly forced at the surface are stationary (i.e. they have zero zonal phase speed) and are therefore not able to propagate into easterly winds. This fact is responsible for the quiet summer stratosphere with Rossby waves decaying within the lower stratosphere. During winter when winds are westerly, Rossby waves are able to propagate into the stratosphere, however there is a cutoff for propagation at stronger westerly wind velocities which decreases to lower wind speeds for increasing horizontal wave numbers. On a β-plane, waves of zonal and meridional wave numbers k and l and phase speed c are able to propagate into uniform background winds u_0 fulfilling the following criterion:

$$0 < u_0 - c < u_c = \frac{\beta}{(k^2 + l^2) + f_0^2/4H^2N^2}$$

(Charney and Drazin, 1961), where u_0 is the zonal mean background wind, u_c is the critical velocity, $\beta = \frac{\partial f}{\partial y}$ around a latitude of $\varphi_0 = 45°$, $f_0 = 2\Omega \sin(\varphi_0)$ is a constant midlatitude value of the Coriolis parameter, H is the scale height and N is the buoyancy frequency. The climatology of the stratospheric winds indicates that according to this criterion only the longest waves (zonal wave numbers 1 and 2) are able to propagate in Northern Hemisphere midwinter, while no waves may be able to penetrate the strong Southern Hemisphere midwinter winds, explaining the differences between the hemispheres outlined above (Plumb, 1989).

It can not, however, be anticipated from the above propagation characteristics that these long waves tend to occasionally amplify within the extratropical atmosphere. They propagate into the upper stratosphere and break, thereby impacting the polar vortex winds. This behavior of the waves often leads to a spectacular and sudden weakening of the stratospheric polar vortex in *stratospheric sudden warming* events. Although the planetary waves which are able to propagate into the stratosphere are dominantly forced at the Earth's surface, lower tropospheric precursor signals are not significantly different from the average state of the troposphere, and a signal of significant wave amplification can often only be found above the tropopause around an atmospheric pressure of 100 hPa (Polvani and Waugh, 2004). This indicates that the stratosphere plays a major role in amplifying or modifying the wave flux entering the stratosphere, particularly in the lead-up to stratospheric warming events.

The observed wave amplification within the stratosphere illuminates the two-way interaction between the zonally averaged flow and the waves described in Section 1.2.1: while the possibility for and the direction of wave propagation as well as the breaking location are determined by the strength and distribution of the mean flow, breaking waves are in turn able to weaken the westerly flow. The question arises if the wave flux entering the stratosphere is able to continuously alter the stratospheric background flow to either allow for an increased wave propagation or to alter the pathway of propagation into the upper stratosphere. A poleward focus of wave energy has been observed before stratospheric warmings [e.g. Limpasuvan et al. (2004)]. Possible explanations for the evolution into sudden warmings involve the idea of *resonant self-*

tuning (Plumb, 1981) and cavity theory (Matsuno, 1970) as described in the next section.

The interaction of the zonally averaged mean wind with the zonal anomalies as well as possibly also the interaction between waves are therefore important for understanding the short term variability of the stratospheric flow as well as possible impacts on the tropospheric flow.

1.3.1 Classification of Sudden Warmings

Sudden warmings are generally classified into wave-1 (*displacement*) and wave-2 (*split*) events (O'Neill, 2003). For wave-1 warmings, the polar vortex is shifted off the pole by a strong anticyclone. In the Northern Hemisphere, a strong Aleutian high forms over the North Pacific in the winter stratosphere [e.g. Limpasuvan et al. (2004)]. The stratospheric Aleutian high corresponds to a strong Aleutian low at the surface due to the phase tilt with height of the stationary wave. For wave-2, a second anticyclone forms opposite to the Aleutian anticyclone; these then act together to weaken and split the vortex.

For warmings classified as displacements it is often observed that a strong wave-1 precursor weakens the stratospheric polar vortex over the weeks before the warming (usually with a maximum about two weeks before the warming), which then culminates in a strong wave-1 burst causing the displacement off the pole [e.g. Labitzke (1977), Labitzke (1979), Labitzke (1981)]. For split warmings, wave-2 is mainly responsible for the burst at the time of the warming, while wave-1 plays a dominant role for the precursor wave activity during the weeks before the warming (Labitzke, 1981). This may seem surprising but can possibly be explained by the necessity to weaken the vortex ahead of time in order for a warming to occur. Since wave-1 is able to propagate into stronger winds than wave-2, it is more likely to be responsible for an initial weakening of the wind. Chapter 3 will give more detailed information about the identification and classification of sudden warmings.

The differences between the rather undisturbed Southern Hemisphere winter stratosphere and the strong variability in the Northern Hemisphere winter stratosphere

translate into the frequency of occurrence of sudden warmings. The Northern Hemisphere exhibits on average about 6 events per decade [e.g. Charlton and Polvani (2007)], divided approximately equally between splittings and displacements. In the Southern Hemisphere, there is evidence for minor warmings [e.g. Labitzke (1981), Quiroz (1966), Labitzke and van Loon (1972)], but since observations have started to be routinely performed by satellites only one major warming has been observed (on September 26, 2002). This warming was characterized as a wave-2 (split) warming. The anomalous winter of 2002 has been widely investigated [e.g. in a special issue by the *Journal of Atmospheric Sciences*: Newman and Nash (2005), Charlton et al. (2005), Kushner and Polvani (2005)], but the exact cause of the anomalous Southern Hemisphere stratospheric warming is still a matter of ongoing research. The fact that the Southern Hemisphere winter polar vortex is only weakly disturbed may lead to the conclusion that only stationary waves are able to cause strong sudden warmings, as observed in the Northern Hemisphere. On the other hand, the causality is not clear: a stronger polar vortex, as observed in the Southern Hemisphere, requires more weakening by the waves for a sudden warming to occur, while on the other hand the strong vortex winds prevent waves from propagating into the upper stratosphere.

In addition, a connection of sudden warmings to the tropical stratosphere has been found, which yields meridional connections beyond the vertical propagation criterion by Charney and Drazin (1961): The strength of the polar vortex is reduced during the easterly phase of the quasi-biennial oscillation (QBO) (Holton and Tan, 1980). The proposed mechanism involves a change in the structure of the wave guide due to the mean wind changes in the tropical lower stratosphere. Several studies have confirmed and expanded this relationship [e.g. Labitzke and van Loon (1988), Gray et al. (2001), Garfinkel et al. (2012)].

1.3.2 Theory and Modeling Approaches to Stratospheric Sudden Warmings

Stratospheric sudden warming - type variability can be simulated in models over a large range of complexities and degrees of freedom from simple mechanistic models to complex non-linear and fully three-dimensional weather prediction models. All of these models are able to capture variability comparable to sudden warmings appropriate to the level of model complexity. This section will give a brief overview of the history and progress in the understanding and modeling of sudden warmings.

Initially, motivated by the sudden nature of the wave amplification and the breakdown of the vortex, instability hypotheses were formulated. Hypotheses of both baroclinic and barotropic instability of the polar vortex were considered. However, Charney and Stern (1962) performed a stability analysis for the polar vortex and concluded that it did not meet the necessary conditions for instability. Dickinson (1973) and Simmons (1974) concluded that unstable modes are not responsible for sudden warmings due to their insufficient growth rates. While the vortex itself does not become unstable, the potential vorticity gradient can be reversed locally when a filament of potential vorticity is drawn out of the vortex, as observed during strong wave events. This mechanism is however not responsible for the sudden warming.

It was recognized relatively early on that sudden warmings are caused by waves propagating up from the troposphere, which amplify in the stratosphere and interact with the mean flow. Matsuno (1971) was able to construct a mechanistic model of a sudden warming by representing only the basic components of the stratosphere which he found to be important for the development of a warming, i.e. a strong wave pulse propagating into the stratosphere as well as the ability for the waves to interact with the mean flow. He treated the troposphere as a source of wave energy while using stratospheric mean winds from data. The model showed that a transient wave pulse is able to weaken the wind above the vortex, and that the signal then propagates down into the stratosphere to induce a stratospheric sudden warming through the interaction of the wave with the critical line, i.e. the place where the phase speed of

the wave matches the wind speed.

Based on the finding that the interaction of the mean wind with the wave is crucial to force a warming, Holton and Mass (1976) constructed a model exhibiting the characteristic features of the extratropical stratosphere: a relaxation of the mean state to radiative equilibrium (on weekly time scales) paired with the interaction of the waves with the mean flow (on daily time scales). Their model was based on a quasi-geostrophic β-plane channel model to study wave-mean flow interaction in the stratosphere. A wave was induced at the model lower boundary by a steady forcing, while the mean state was relaxed to an equilibrium profile using Newtonian cooling. They found that for a low amplitude wave forcing, the stratospheric circulation remained steady and close to radiative equilibrium, while for a larger forcing amplitude the response in the waves and the mean flow exhibited quasi-periodic oscillations. This important result has been reproduced and extended in various studies [e.g. Yoden (1987)]. For additional details about the early modeling approaches for sudden warmings see e.g. Schoeberl (1978).

While these early models illustrate the importance of stratospheric internal variability for sudden warmings, the reason as to why the waves amplify in the first place to cause sudden warmings is still unclear. The nature of the sudden amplification inspired theories of resonant states of the stratosphere - troposphere system. Matsuno (1970) proposed a cavity theory, where waves are trapped within a leaky cavity in the extratropical atmosphere, bounded by a weak potential vorticity gradient equatorward, yielding a wave guide into the upper stratosphere, and reflective layers on top of the cavity set up by the mean wind structure. It is found that waves tend to deviate from their climatological tendency to propagate equatorward to focus into the polar regions before sudden warmings [e.g. Limpasuvan et al. (2004)], while in addition reflective surfaces have also been found in the lower stratosphere [e.g. Harnik and Lindzen (2001), Perlwitz and Harnik (2003)], which tend to strengthen the wave flux below the reflective layer. The sudden nature of the warmings could then be explained by a sudden leak in the reflective layer allowing the wave energy stored within the cavity to propagate into the upper stratosphere.

While baroclinic and barotropic instabilities were largely ruled out as a direct cause for sudden warmings, Plumb (1981) found that another type of unstable interaction may contribute to the evolution into sudden warmings: the interaction and positive feedback between free internal and forced stationary waves may lead to an amplification as the free mode slows down to match the phase speed of the stationary wave, leading to a *self-tuned resonance*. Events of traveling waves slowing down have been observed in the stratosphere before sudden warmings. Recently, this theory has been revisited and applied to both splitting (Esler et al., 2006) and displacement events (Esler and Matthewman, 2011).

The above theories for wave amplification may be able to explain both the sudden burst in wave energy as well as the slow buildup of wave energy entering the stratosphere. Precursor structures and the accumulation of wave energy have been shown to be crucial for the evolution into a sudden warming. So far the only indicator for predicting the strength of the stratospheric flow seems to be given by the wave flux entering the stratosphere integrated over several weeks (Polvani and Waugh, 2004).

One possible shortcoming of many of the described models is the inability of the waves to propagate equatorward, and they have been criticized for their rigid boundaries on the equatorward side. Climatologically, waves tend to propagate up and equatorward within the troposphere, while during winter waves are also able to propagate into the stratosphere. Rigid latitudinal boundaries inhibit the waves from following their climatological equatorward path, which yields a model atmosphere in which the waves are constantly focused upward, possibly allowing for more sudden warmings to occur. These models are therefore not able to distinguish between the observed focusing of wave energy into the polar regions as opposed to equatorward propagation in order to reproduce a realistic frequency of sudden warming events. The possibility for the waves to propagate in the meridional direction is important for a realistic atmosphere and for reproducing sudden warmings, as will be shown in Chapter 5. However, it can be shown that even when allowing for meridional wave propagation (Scott and Polvani, 2004), the stratospheric vacillation cycles can be reproduced. The model atmosphere studied in Scott and Polvani (2004) is in principle

similar to the one in Holton and Mass (1976), but allows for latitudinal propagation of the waves. In addition, their model includes wave-wave interaction in the stratosphere, however this does not seem to strongly impact the forcing of stratospheric warmings in their model. Similar complex models allowing for stratospheric variability only [e.g. Scott and Polvani (2006)] show that stratospheric internal variability by itself is able to reproduce stratospheric vacillation cycles.

Weather forecasting models routinely forecast variations in the stratosphere in addition to tropospheric weather. In addition, weather forecasting models focusing on stratospheric variability are developed in parallel with higher model lids and increased stratospheric resolution [e.g. Bancalá et al. (2012)]. However, despite the much slower time scales dominant in the stratosphere as compared to the troposphere, sudden warmings are difficult to be predicted with current weather models. For dynamical core models, Gerber et al. (2009) show in an ensemble study that sudden warmings can be predicted given stratospheric initial conditions 10 days in advance, while the ability to forecast, as expected, decreases with increased lead time, with a threshold value around 20 days for their choice of initial perturbation beyond which the warmings are not predicted by the ensemble.

1.4 Stratosphere - Troposphere Coupling

1.4.1 Stratosphere - Troposphere Coupling during Stratospheric Sudden Warmings

Since the waves responsible for causing stratospheric sudden warmings enter the stratosphere from the troposphere, it makes sense to look for tropospheric structures which resemble the stratospheric waves in the lead-up to stratospheric warmings. Stationary Rossby waves are mainly caused by surface topography and heating patterns in the lower troposphere. Large-scale Rossby waves may in addition be caused by baroclinic instability in the upper troposphere (see Chapter 4), however for these there is no reason for stationarity. The stratosphere exhibits both traveling and sta-

tionary Rossby wave patterns, with traveling waves more important in the Southern Hemisphere with respect to the Northern Hemisphere due to the weaker stationary wave forcing.

However, it is very difficult to link stratospheric variability to tropospheric wave structures and it is currently not (and possibly never will be) possible to predict sudden warmings from tropospheric precursors due to the strong internal variability of the troposphere. It can however be observed that planetary-scale wave anomalies at the tropopause are clearly linked to stratospheric wave propagation. For this reason, various studies have looked into the connection between tropospheric and stratospheric wave-1 and wave-2 patterns.

Atmospheric blocking has often been invoked as a possible connection between tropospheric variability and sudden warmings. The stratosphere has been shown to require a large-scale, long-lived and deeply penetrating forcing, and blocking patterns are among the longest-lived extratropical planetary-scale tropospheric atmospheric structures, which sometimes even exert influence into the lower stratosphere. Martius et al. (2009) show that wave-1 patterns exhibit consistent longitudinal structures which align with the sudden warming pattern before the warming occurs. Wave-2 also exhibits a consistent longitudinal structure but is not always able to grow into the stratosphere. These patterns tend to align with areas of major blocking frequency in the tropopause region.

Various other studies have tried to show a connection between blocking and sudden warmings, but causality is difficult to show and would have to be investigated in a detailed model study with an explicit cause and effect relationship, or with a reanalysis or data study computing the significance of blocks before a warming with respect to blocking in the absence of sudden warming events. The example of blocking illustrates how even a strong tropopause signal is not unambiguously linked to stratospheric variability.

While the connection between tropospheric precursors and sudden warmings is debated, a connection to surface climate has been observed and modeled *after* the occurrence of a sudden warming (Baldwin and Dunkerton, 2001). While model studies

[e.g. Polvani and Kushner (2002), Baldwin and Dunkerton (2001)] show that an increase in the strength of the polar vortex induces a poleward shift of the tropospheric jet, the exact mechanism of stratosphere - troposphere coupling is only partly understood. Proposed pathways include internal wave reflection of extratropical Rossby waves (Harnik and Lindzen, 2001), *downward control* through a meridional circulation induced by the wave forcing (Haynes et al., 1991) and local wave - mean flow interaction [Matsuno (1971), Plumb and Semeniuk (2003)]. These theories will in more detail be explored in Chapter 5.

1.4.2 Stratosphere - Troposphere Coupling in the Winter Hemisphere

A similar connection between the strength of the polar vortex and the latitudinal location of the tropospheric jet can be observed in long term studies. Observational analysis has shown a poleward shift of the Southern Hemisphere tropospheric jet over the past 30 years (Thompson et al., 2003), which was shown in model studies to be dominantly caused by ozone depletion over Antarctica [e.g. Gillett and Thompson (2003), Son et al. (2008), Polvani et al. (2011)]. An increase in greenhouse gases has been shown to have the same-signed effect on the location of the tropospheric jet (Kushner et al., 2001). In the Northern Hemisphere, projections show that a warming of the tropical troposphere along with a radiative cooling of the polar stratosphere through increasing CO_2 strengthens the lower stratospheric meridional temperature gradient, which may be able to drive a poleward shift of the eddy-driven jet (Chen and Held, 2007).

The observed connection between the tropospheric jet and the stratospheric polar vortex can be framed in terms of the annular modes, which will be used as an indicator of downward influence in Chapter 5. These modes represent the dominant patterns of climate variability in the mid-latitudes and polar regions of both hemispheres on weekly to monthly time scales [e.g. Thompson and Wallace (1998), Thompson and Wallace (2000)]. As expected from the variability described above,

the annular mode corresponds to a meridional shift of the tropospheric jet, while in the stratosphere, the mode represents the variability in the strength of the polar vortex, and the stratospheric and tropospheric patterns exhibit a strong connection during times of strong stratospheric variability in Northern Hemisphere winter and in Southern Hemisphere spring (Thompson and Wallace, 2000). The patterns can be observed in the surface signature of the various meteorological fields: geopotential, surface pressure as well as the zonal wind field [Lorenz and Hartmann (2001), Lorenz and Hartmann (2003)]. Annular mode variability is therefore suggested to be closely related to surface anomalies in temperature and pressure in the Northern Hemisphere [e.g. Thompson and Wallace (2000), Baldwin (2001)] and over Antarctica (Kwok and Comiso, 2002), indicating a possible pathway between stratospheric variability and surface climate.

The annular modes appear as a response to external forcing as well as natural internal variability [Ring and Plumb (2007), Ring and Plumb (2008)]. The exact definition of the annular modes is still debated and is a reason for incoherent results comparing different studies (for a summary and evaluation see e.g. Baldwin and Thompson (2009) or Thompson et al. (2003)). Including the stratosphere further complicates the definition due to the strong temporal variability of stratosphere - troposphere coupling and the inclusion of stratospheric layers which account for a comparably small amount of mass in the atmosphere. The decreasing density with height is usually dealt with by weighting the atmosphere according to the mass a level accounts for (Thompson and Wallace, 2000). In addition, the decrease in area with latitude is adjusted for. However, the annular mode patterns are dependent on the applied weighting as will be shown in Chapter 6.

1.5 Terminology

The term *mean flow* will apply to the zonally averaged state of the zonal wind. Correspondingly, the term *wave* will apply to deviations from the zonally averaged flow, i.e. wave-1 corresponds to a zonal anomaly of zonal wave number 1.

Stratospheric Sudden Warmings will be described by the terms *stratospheric sudden warmings*, *sudden warmings*, *stratospheric warmings*, but also, if a confusion can be ruled out, simply *warmings*.

1.6 Thesis outline

This thesis addresses the coupling mechanisms between the stratosphere and the troposphere. Since this coupling is particularly strong during the lead-up to stratospheric sudden warmings as well as in the aftermath of these warmings, the coupling is investigated during the winter season, when tropospheric planetary-scale waves are able to propagate into the stratosphere and cause significant variability in the stratosphere in terms of minor or major sudden warmings. The analysis will address both the coupling brought about by the upward propagating waves as well as the pathway leading to a possible tropospheric response to a stratospheric forcing.

Chapter 2 will introduce the model used for this study and the model runs performed in order to investigate the above questions. The model includes the ingredients which were identified as necessary for a sudden warming: Sufficient stratospheric resolution and a high top are used for the model, which has been shown to be sufficient to cause stratospheric warmings in Gerber and Polvani (2009). Temperature is relaxed to a zonally symmetric equilibrium temperature profile, mimicking the behavior of the mean flow relaxing to radiative equilibrium. In addition, a wave forcing is applied in the troposphere in several of the runs to force stratospheric variability. The type of forcing is the main parameter being varied between the different model runs. Latitudinal and longitudinal propagation are possible in the three dimensional model to allow for the characteristic equatorward propagation of the waves as well as to allow for longitudinally traveling waves which dominate the Southern Hemisphere. Chapter 3 then investigates the sudden warmings obtained in the model run and compares them to reanalysis data in order to ensure that the model warmings exhibit the features which are expected to be reproduced.

Using the above tools, the remaining chapters address stratosphere-troposphere

coupling during both strong and weak stratospheric variability with the goal of characterizing the mechanisms responsible for the vertical coupling. An important part of the upward coupling is given by wave propagation into the winter stratosphere. In the Northern Hemisphere, the forcing and propagation of these waves is well understood, while the Southern Hemisphere is characterized by traveling waves whose generation in the troposphere is not fully understood. Chapter 4 examines the nature of traveling waves, both in terms of their role in the stratosphere as well as in terms of the tropospheric generation mechanism. A second important part of the vertical coupling is the exchange of momentum, which happens both through wave propagation as well as through other mechanisms which may be able to induce a tropospheric response following a strong anomaly in the stratospheric flow. Chapter 5 analyzes the nature of the wave flux into the stratosphere before the warming in terms of timing, location and strength in order to improve the understanding of the momentum balance. In addition, annular mode variability is examined in the angular momentum framework in order to illuminate the tropospheric response to the stratospheric anomaly. While Chapters 4 and 5 are not able to give a concluding answer on the vertical coupling mechanism, they both hint at the importance of synoptic eddy variability in the troposphere for both the forcing of tropospheric traveling waves as well as for the downward influence of stratospheric signals. Since the effects of baroclinic eddies are difficult to parameterize in the model, Chapter 6 revisits the coupling mechanism between the stratosphere and the troposphere during sudden warming events using statistical analysis. Synoptic-scale eddies are able to induce an annular mode response which is generally represented by Empirical Orthogonal Functions (EOFs). However, EOF patterns are dependent on weighting structures and therefore may not be the ideal method to consider for stratosphere - troposphere coupling. Principal Oscillation Pattern (POP) analysis is applied in an attempt to improve the characterization of the annular modes and to find the physical modes of the stratosphere - troposphere system. The last chapter will offer conclusions summarizing and discussing the results and implications of the findings.

Chapter 2

Description of the Model and the Performed Model Runs

2.1 Model Setup

The model prominently used throughout this thesis is the three dimensional spectral core of a general circulation model of intermediate complexity, comparable to the model used in Scott and Polvani (2004) and Gerber et al. (2009). The code used is the *Jakarta* version of the Geophysical Fluid Dynamics Laboratory (GFDL) Flexible Modeling System (FMS) spectral dynamical core. The model setup is largely based on the model description in the appendix of Polvani and Kushner (2002).

The model runs have no seasonal cycle and are run in perpetual mid-winter conditions in order to allow for large statistics of mid-winter sudden warmings, excluding final warmings which are induced by the polar vortex weakening in Spring due to changes in the radiative forcing, helped along by planetary waves. For mid-winter warmings, the forcing is solely provided by the waves, which is precisely the mechanism we would like to study.

Waves are able to propagate in three dimensions, and wave-wave interaction is allowed by the model setup. The model runs performed for this thesis will investigate the stratospheric and tropospheric processes by employing different forcing mechanisms: the control run will not force stationary waves, thereby simulating the flow

in the Southern Hemisphere, while the topography run simulates a Northern Hemisphere winter climatology. The truncated run illuminates the processes present in an atmosphere inhibiting synoptic waves as well as the role of free traveling waves. This chapter gives a more detailed description of the model setup, the performed runs and how sudden warmings are modeled in this particular study.

2.1.1 Spectral Setup

The model uses a spectral setup, thereby simplifying a possible truncation in wave number space as employed later in this study (Section 2.2.3). All presented model runs were performed at a spectral truncation of T42, corresponding to a $2.8^\circ \times 2.8^\circ$ horizontal resolution on a latitude / longitude grid. There are several reasons for this choice of resolution: Previous model studies have explored the parameter range for model resolution, and e.g. Gerber and Polvani (2009) found that T42 and T85 exhibit a very similar atmospheric variability, and both produce warming events about every 200 days in the topography run presented in Section 2.2.2. A resolution of T21 has consistently been shown to not be sufficient for the study of sudden warmings [e.g. Polvani and Kushner (2002), Gerber and Polvani (2009)]. A resolution of T85 however turned out to be too costly to perform long model runs to yield good warming statistics. As a sensitivity study, a topography model run with T30 resolution was performed, exhibiting a comparable climatology of sudden warmings comparable to T42. However, for the present study only the T42 run was used in order to facilitate the comparison of the results with the current literature which tends to use T42 for comparable idealized dynamical core model studies. A hyperdiffusion of ∇^6 is used for all model runs to damp small-scale structures.

2.1.2 Hybrid Vertical Coordinates

In the vertical, 40 levels were used between 925 hPa and 0.02 hPa in hybrid coordinates. Hybrid coordinates correspond to a topography-following σ coordinate in the troposphere with a smooth transition to pure pressure coordinates in the strato-

sphere. The hybrid vertical differencing is given by Simmons and Burridge (1981). Hybrid vertical levels are generally used for numerical spectral core models resolving the stratosphere [e.g. the WACCM model in Garfinkel and Hartmann (2010), and MPI ECHAM3, UKMO HadAM1, NCAR CCM2]. While earlier versions of the here used model with a similar setup use pure sigma coordinates [e.g. Gerber and Polvani (2009), Polvani and Kushner (2002)], it is here found that using sigma coordinates in the present model yields the same problem as described in Trenberth and Stepaniak (2002) for the NCEP reanalysis (Kalnay et al., 1996), which suggests the use of hybrid coordinates. A sensitivity study using pure σ coordinates was compared to the results using hybrid coordinates and no difference was found except within the top model layers (within the sponge layer described in Section 2.1.4) which were not used for the analysis of the model results.

This description will now broadly follow the official FMS model description of the spectral core. Pressure at the mid-levels of the vertical grid is given by

$$p_{k-1/2} = A_{k-1/2}\, p_{ref} + B_{k-1/2}\, p_s \tag{2.1}$$

for $k = 1, N+1$, where p_{ref} is a constant reference pressure, here chosen to be 1000 hPa, close to the global mean surface pressure. It ensures that A can be chosen to be dimensionless with values between 0 at the surface and 1 at the top of the model. For $A(k) = 0 \; \forall \; k$, the coordinate system reduces to pure σ coordinates, while $B(k)$ determines the level spacing. For $B(k) = 0 \; \forall \; k$, the coordinate system reduces to pure pressure coordinates.

For a hybrid coordinate system, the σ coordinate for the bottom layers can then be defined using $[A_\sigma(k) = 0, B_\sigma(k)]$, and the pressure coordinates for the top layers can be defined using $[A_p(k) = B_\sigma(k), B_p(k) = 0]$. The transition between the σ and pressure coordinates is given by

$$A(k) = \gamma(k)A_\sigma(k) + [1-\gamma(k)]A_p(k) = [1-\gamma(k)]B_\sigma(k) \quad (2.2)$$

$$B(k) = \gamma(k)B_\sigma(k) + [1-\gamma(k)]B_p(k) = \gamma(k)B_\sigma(k) \quad (2.3)$$

with $\gamma(k) = 1$ near the surface and $\gamma(k) = 0$ near the model top.

The model output is interpolated onto pressure levels throughout the model domain for the performed analyses, i.e. all plots shown in this thesis are given on pressure levels, except where indicated.

2.1.3 Relaxation to Equilibrium Temperature Profile

The model employs Newtonian cooling specified by a relaxation to a specified temperature profile (Held and Suarez, 1994), which is adjusted to control the strength of the polar vortex and the location of the tropospheric jets, respectively. The equilibrium temperature profile is computed separately in the stratosphere and in the troposphere. In the stratosphere it is given by

$$T_{eq}^{strat}(p,\varphi) = [1-W(\varphi)]T_{US}(p) + W(\varphi)T_{PV}(p), \quad (2.4)$$

where φ is latitude and p is pressure, T_{US} is the *US Standard Temperature (1976)*, and

$$T_{PV}(p) = T_{US}(p_T)\Big(\frac{p}{p_T}\Big)^{R\gamma/g} \quad (2.5)$$

(compare to Polvani and Kushner (2002): note the sign error in the equation given in their appendix) and with the weight function

$$W(\varphi) = 0.5[1-\tanh((\varphi-\varphi_0)/\delta\varphi)] \quad (2.6)$$

with $\varphi_0 = -50°$ and $\delta\varphi = 10°$. The parameter γ is used to determine the strength of the stratospheric polar vortex by changing the lapse rate of the equilibrium temperature profile T_{eq}. Typical values are between $\gamma = 1\,\mathrm{K/km}$ and $\gamma = 4\,\mathrm{K/km}$, with

$\gamma = 4\,\mathrm{K/km}$ corresponding to a strong Southern Hemisphere - like vortex, which was used throughout this thesis. A case with no stratospheric vortex is obtained by setting the weight function $W(\varphi) = 0$.

The equilibrium temperature profile for the troposphere is given by

$$T_{eq}^{trop}(p,\varphi) = max[T_T, (T_0 - \delta T)\Big(\frac{p}{p_0}\Big)^{\kappa}], \tag{2.7}$$

where $T_0 = 315\,\mathrm{K}$, $p_0 = 1000\,\mathrm{mb}$, and $\kappa = 2/7$, with

$$\delta T = \delta_y \sin^2(\varphi) + \epsilon \sin(\varphi) + \delta_z \log(p/p_0)\cos^2(\varphi), \tag{2.8}$$

where $\delta_y = 60\,\mathrm{K}$, $\delta_z = 10\,\mathrm{K}$. Continuity between the stratospheric and the tropospheric equilibrium temperature profiles at $p = p_T$ is given by $T_T = T_{US}(p_T)$. Figure 2-1 gives the radiative equilibrium temperature T_{eq} for the winter hemisphere in degrees Kelvin.

The parameter ϵ is used to describe an asymmetry between the summer and winter hemispheres in order to simulate winter or summer conditions in the respective hemispheres, where $\epsilon = 0\,\mathrm{K}$ corresponds to equinox. Increasing ϵ corresponds to an increasing asymmetry between the hemispheres, with positive (negative) values of ϵ describing Northern (Southern) Hemisphere winter. In this study, a value of $\epsilon = -10\,\mathrm{K}$ was used in order to yield Southern Hemisphere winter conditions.

The relaxation to the equilibrium temperature profile is given by

$$\frac{\partial T}{\partial t} = -s_{damp}(T - T_{eq}) \tag{2.9}$$

where s_{damp} is the damping parameter in units of days^{-1} given by

$$s_{damp} = \begin{cases} \frac{1}{k_a} + \cos^4(\varphi)\,\frac{1/k_s - 1/k_a}{\sigma - \sigma_b} & \text{below } \sigma_b \\ \frac{1}{k_a} & \text{above } \sigma_b \end{cases} \tag{2.10}$$

where $k_s = 4\,\mathrm{days}$, $k_a = 40\,\mathrm{days}$, $\sigma_b = 0.7$. The relaxation timescales are illustrated in Figure 2-2.

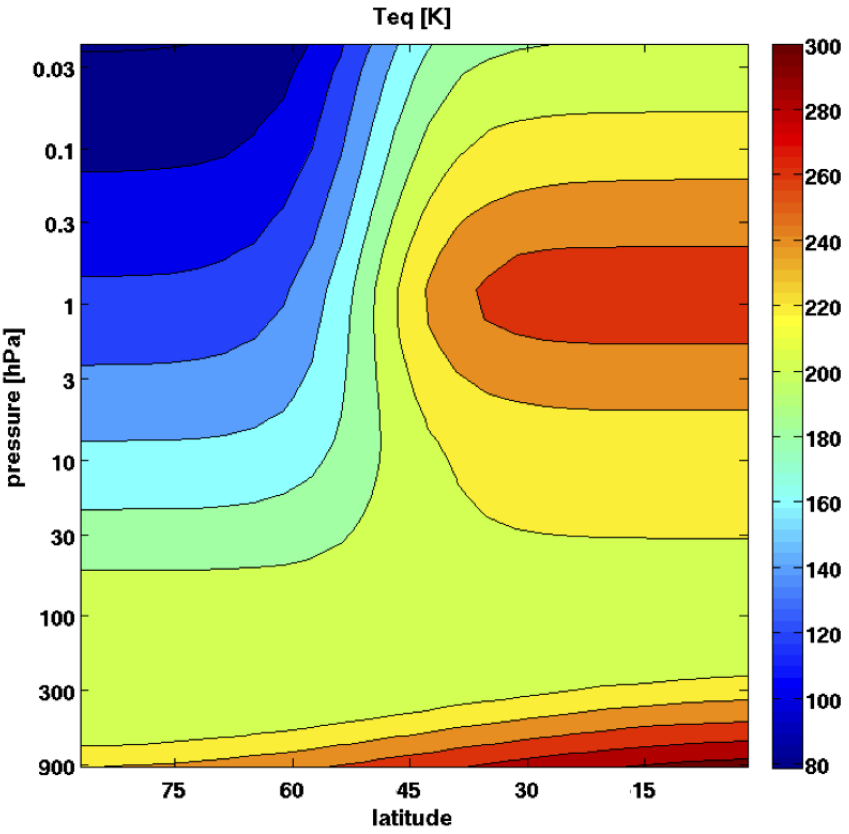

Figure 2-1: Zonally symmetric equilibrium temperature profile for the model run in [K] for the winter hemisphere as a function of latitude and pressure. Contour interval: 20 K.

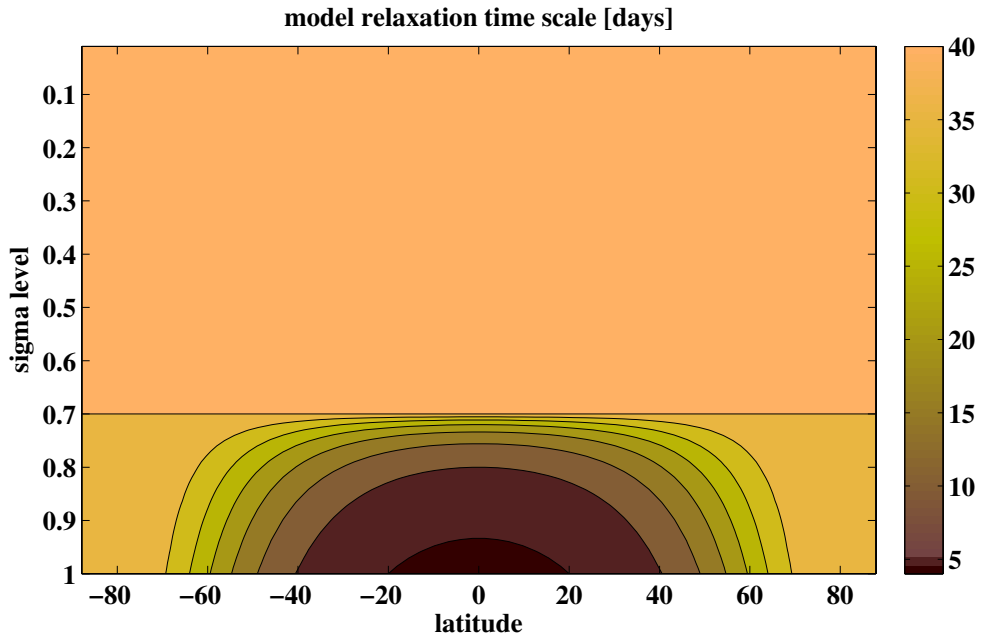

Figure 2-2: Time scales for the relaxation to equilibrium temperature in units of days given by s_{damp}^{-1}. Contour interval: 5 days.

2.1.4 Damping at the Model Top

A sponge layer is applied within the top layers of the model in order to prevent artificial wave reflection from the top of the model. These levels are therefore not used for any of the subsequent analysis. The horizontally symmetric sponge layer is given by a linear damping on the $x-$ and $y-$momentum equations, with a damping coefficient given by

$$k_{sp}(p) = \begin{cases} k_{max}\left(\frac{p_{sp}-p}{p_{sp}}\right)^2 & \text{for } p < p_{sp} \\ 0 & \text{for } p \geq p_{sp} \end{cases} \tag{2.11}$$

with $p_{sp} = 0.5\,\text{mb}$ and $k_{max} = 2\,\text{day}^{-1}$ (note that this is a corrected version of the number given in the appendix of Polvani and Kushner (2002)). The damping is then applied using

$$\frac{\partial u}{\partial t} \rightarrow \frac{\partial u}{\partial t} + k_{sp} \cdot u \tag{2.12}$$

$$\frac{\partial v}{\partial t} \rightarrow \frac{\partial v}{\partial t} + k_{sp} \cdot v. \tag{2.13}$$

2.1.5 Damping at the Model Surface

In order to account for the damping of the winds near the surface, Rayleigh damping of the horizontal wind components was employed below the $\sigma_b = 0.7$ level given by

$$v_{damp} = \begin{cases} -\frac{1}{k_f} \cdot \frac{\sigma - \sigma_b}{1.0 - \sigma_b} & \text{below } \sigma_b \\ 0 & \text{above } \sigma_b \end{cases} \tag{2.14}$$

$$\frac{\partial u}{\partial t} = v_{damp} \cdot u \tag{2.15}$$

$$\frac{\partial v}{\partial t} = v_{damp} \cdot v \tag{2.16}$$

where $k_f = 1\,\text{day}$. Note that v_{damp} is zero or negative.

2.2 Performed Model Runs

Different model runs were implemented in this thesis with a focus on yielding different regimes of stratospheric and tropospheric variability. The main runs were run out to 30,000 days (including a spin-up period of 100 days which was removed for the analysis). The remaining model runs have a length of 10,000 days each. The specifics of the most important model runs are summarized in Table 2.1. The main difference between the performed runs is the method by which planetary-scale Rossby waves are forced.

2.2.1 Control Run

A control run was performed employing the specifications described above without topography. This run is designed to be used as a control run to compare to the subsequent runs. It comes closest to a Southern Hemisphere climatology in terms of exhibiting no forcing asymmetry in the zonal direction. It exhibits traveling planetary wave patterns in the stratosphere (see Chapter 4 for a more detailed discussion of this run). Figure 2-3 shows the zonal mean zonal wind averaged over the control run.

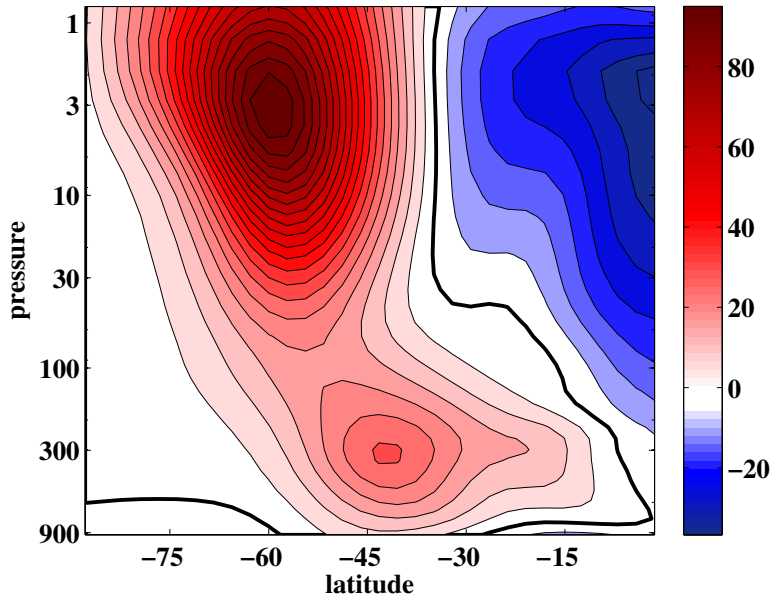

Figure 2-3: Control run: zonal mean zonal wind in ms^{-1}, the contour interval is $5\,\mathrm{ms}^{-1}$. The zero wind line is printed in bold, westerly winds are shaded in red, easterly winds are shaded in blue. The wind is averaged over the first 8500 days of the run, excluding a 100 day spin-up period.

Despite the absence of a longitudinally varying forcing in this run and in other runs with weak or no topography, stratospheric variability and even sudden warming events can arise in these runs. While no sudden warmings were observed in the present run, the amount of stratospheric variability is very variable throughout the run, i.e. very quiet periods are interspersed with periods of stronger stratospheric variability. Several studies have found infrequent but strong sudden warmings for weak or no topography [Gerber and Polvani (2009), Scinocca and Haynes (1998), Kushner and Polvani (2005)], suggesting that a much longer version of the present control run may in fact exhibit sudden warmings.

2.2.2 Topography Run

The goal of the topography run was to achieve a run that exhibits major sudden warmings in a similar fashion to the real atmosphere, in order to do various kinds of statistics. The topography run largely follows *run 9* from Gerber and Polvani

(2009), with the only differences being the hybrid coordinate system and the Southern Hemisphere being the winter hemisphere in this run. They term this run their *most realistic* run in terms of both the "frequency and structure of sudden warming events" (Gerber and Polvani (2009), p.1928). This run was used here with the purpose of a comparatively realistic representation of the real atmosphere. Topography is specified as described below. The run exhibits the same distribution and variability of zonal mean wind as well as annular mode variability as the run from Gerber and Polvani (2009).

Implementation of Topography

Topography was added as a forcing for planetary-scale Rossby waves. The specific zonal wave-2 mountain implemented was adapted from Gerber and Polvani (2009) by setting the surface geopotential height to

$$Z(\lambda,\varphi) = \begin{cases} h_0 \sin^2\left(\frac{\varphi-\varphi_0}{\varphi_1-\varphi_0}\cdot\pi\right)\cos(m\lambda) & \text{for } \varphi_0 < \varphi < \varphi_1 \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

(note that this is a corrected version of the equation given in Gerber and Polvani (2009)) with λ and φ corresponding to longitude and latitude and values of $m = 2$, $h_0 = 3000\,\mathrm{m}$, $\varphi_0 = -25°$ and $\varphi_1 = -65°$, centering the topography at 45°S.

Wave-2 topography was chosen over wave-1 topography because Gerber and Polvani (2009) found that wave-1 forcing in this model does not produce a realistic climatology of warmings, while instead slowly eroding the vortex and stripping away filaments of potential vorticity. In addition, they found that the regime which produces warmings narrows with decreasing wave number, where wave-1 topography at a height of 4000 m produces a strong vortex with an unrealistically small number of warmings, while topography at 4500 m significantly weakens the vortex without any clear relaxation to radiative equilibrium in between, and no clear warming regime could be observed between these values (Gerber and Polvani, 2009). Considering that the real atmosphere exhibits both strong stratospheric wave-1 and wave-2 forcing, it is not surprising that the model atmosphere does not reproduce a realistic climatology

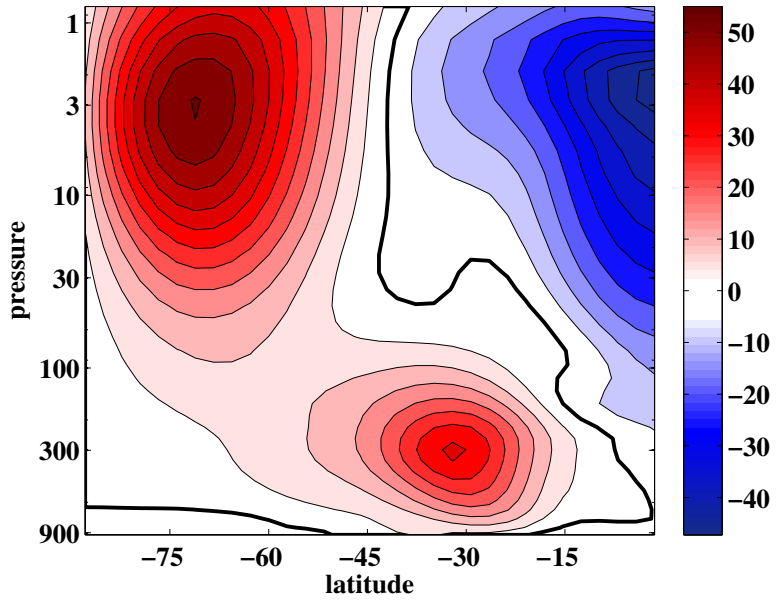

Figure 2-4: Topography run: zonal mean zonal wind in ms^{-1}, the contour interval is $5\,\mathrm{ms}^{-1}$. The zero wind line is printed in bold, westerly winds are shaded in red, easterly winds are shaded in blue. The wind is averaged over the entire run, excluding a 100 day spin-up period.

with only wave-1 forcing. For the same reasons, it is not self-evident that wave-2 forcing would produce a better climatology, but as was shown in Gerber and Polvani (2009), the above forcing parameters do yield a realistic climatology in terms of the frequency, magnitude and evolution of sudden warmings. The reason for not forcing both wave numbers lies in the simplicity which was sought in the model, as well as the limitations of the model described above.

Comparing Figures 2-3 and 2-4 shows that adding topography considerably changes the mean wind structure of both the troposphere and the stratosphere: Due to the stronger wave forcing, more planetary-scale Rossby waves are produced in the topography run. These waves are able to propagate into the stratosphere and they significantly weaken the time averaged polar vortex for the run including topography: The polar vortex in the control run exhibits time mean wind speeds of about $100\,\mathrm{ms}^{-1}$ (Figure 2-3) with a standard deviation of $5\,\mathrm{ms}^{-1}$, while the run including topography exhibits vortex wind speeds of around $57\,\mathrm{ms}^{-1}$ (Figure 2-4) with a standard devia-

tion of $15\,\mathrm{ms}^{-1}$. The variability in the latitudinal location of the tropospheric jet is reduced in the topography run as described in Chan and Plumb (2009) and Gerber and Polvani (2009). The characteristic equatorward shift of the tropospheric jet for a weaker polar vortex as described in Section 1.4.2 is evident from comparing Figures 2-3 and 2-4.

In addition, Chan and Plumb (2009) have found parameter regimes in the present model yielding unrealistically strong coupling between the stratosphere and the troposphere. They found this behavior for model runs with a co-existence of a tropospheric eddy-driven and a subtropical jet, and attributed it to the observed decorrelation times which were too large by up to an order of magnitude for the specific model runs. The parameter choices used for the present model runs ($\epsilon = -10\,\mathrm{K}$, $\gamma = 4\,\mathrm{K/km}$) yield realistic decorrelation times based on the analysis in Chan and Plumb (2009).

2.2.3 Truncated Runs

For the truncated runs, the goal is to reduce the propagating waves to one single or several planetary wave numbers. The idea is that a truncation of the synoptic eddies will reveal the state of the atmosphere in the presence of planetary Rossby waves only, without the influence of synoptic eddy feedbacks in the troposphere. These runs are cleaner to investigate, since planetary wave numbers which are able to penetrate the stratosphere cannot be produced from the interaction of synoptic eddies as described in Scinocca and Haynes (1998), but only from mechanisms involving planetary-scale waves. In addition, for truncating to a single wave number only, the interaction between waves of different wave numbers can be inhibited. As will be described in Chapter 4, allowing only a single wave number to propagate can nevertheless lead to baroclinic instability of this particular wave number, including sufficient amplification in the stratosphere to cause sudden warmings.

Two truncated runs will be presented in this thesis: a run truncated to wave-2 only (*trunc2*), and a run truncated to both wave-1 and wave-2 (*trunc12*). No topography is used for the truncated runs. Figures 2-5 and 2-6 show the zonal mean zonal wind for

these runs. The mean state of the atmosphere is changed considerably as compared to the control and topography runs.

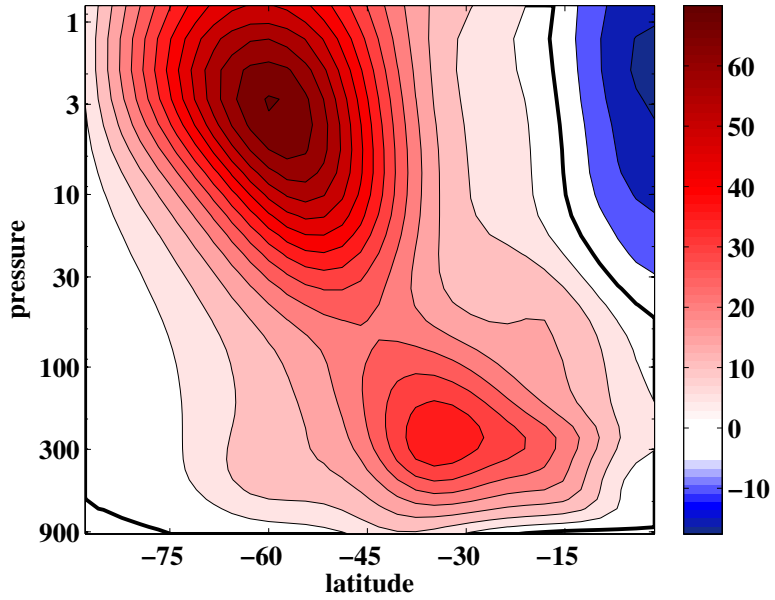

Figure 2-5: Run truncated to zonal wave-1 and wave-2: zonal mean zonal wind in ms^{-1}, the contour interval is $5\,\mathrm{ms}^{-1}$. The zero wind line is printed in bold, westerly winds are shaded in red, easterly winds are shaded in blue. The wind is averaged over the entire run, excluding a 100 day spin-up period.

A more detailed analysis comparing the mean state of the truncated runs to the control run will be shown in Chapter 4.

Implementation of Truncation

The truncation of specific wave numbers can efficiently be implemented in the model due to its spectral setup. The variables which are integrated in time throughout the model run are set to zero at every model time step, except for the spectral components corresponding to the mean flow and the desired wave number(s), e.g. all wave numbers except for zonal wave-2 and the zonal mean flow are set to zero at every time step for the model run truncated to zonal wave-2 (*trunc2*).

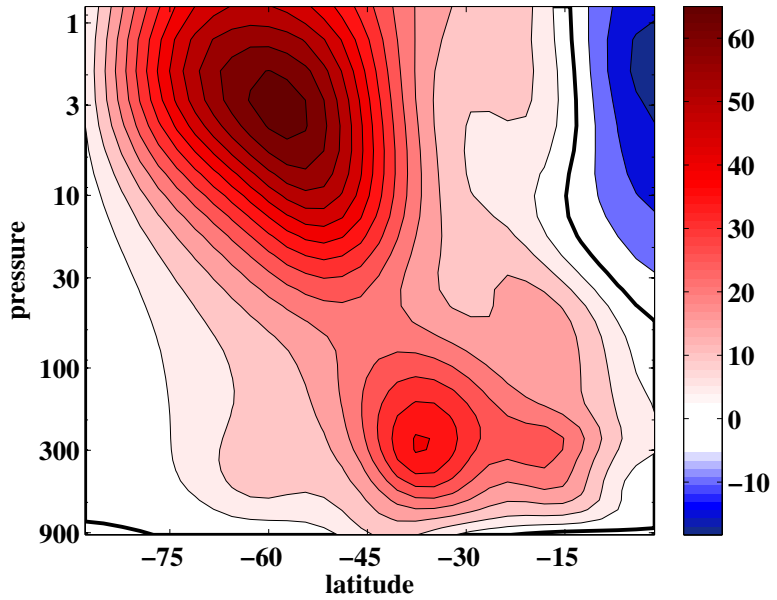

Figure 2-6: Run truncated to zonal wave-2: zonal mean zonal wind in ms^{-1}, the contour interval is $5\,\mathrm{ms}^{-1}$. The zero wind line is printed in bold, westerly winds are shaded in red, easterly winds are shaded in blue. The wind is averaged over the entire run, excluding a 100 day spin-up period.

Surface Damping for the Truncated Runs

The surface damping described in Section 2.1.5 was modified in the truncated runs in order to force westerly winds in the midlatitude lower troposphere to ensure the possibility for wave propagation. Equation (2.15) is modified to

$$\frac{\partial \overline{u}}{\partial t} = v_{damp}\Big[\overline{u} - \sin^2(2\varphi)\, u_{surf}\Big] \tag{2.18}$$

where v_{damp} is negative as given by Equation (2.14) and $u_{surf} = 5\,\mathrm{ms}^{-1}$.

2.2.4 Truncated Runs with Wave Damping

As the truncated runs described above yield significant stratospheric variability as will be described in Chapter 4, experiments with additional wave damping were performed. In order to damp the waves separately from the mean flow, Equation (2.9)

Table 2.1: Statistics for the model runs most prominently used in this thesis from top to bottom: The control run (Section 2.2.1), the topography run (Section 2.2.2), the run truncated to wave-1 and wave-2 as well as the run truncated to wave-2 only (Section 2.2.3). The first column gives the name of the model run, the second column gives the length of the model run excluding a 100 day spin-up time, while columns 3 and 4 give the number of stratospheric sudden warmings detected according to different criteria: Column 3: A sudden warming is detected for the day when the principal component time series of the 10 hPa level computed from zonal mean zonal wind falls below -2 standard deviations (PC2), and no warming is detected for at least 45 days after a warming is identified. Column 4: Same as column 3 but for -3 standard deviations (PC3). Column 5 gives the average frequency of warmings for the criterion in column 3 (PC2). These criteria will be explained in more detail in Section 3.3.

run	length [days]	PC2 warmings	PC3 warmings	PC2 warming frequency
control	9800	-	-	-
topography	29800	154	55	1/194 days
trunc12	9800	28	6	1/350 days
trunc2	29800	106	31	1/282 days

has to be modified to yield

$$\frac{dT}{dt} = -s_{damp}(\overline{T} - T_{eq}) - s_{damp}^{spec}(T - \overline{T}) \tag{2.19}$$

where $\overline{T}$ is zonal mean temperature, T_{eq} is the equilibrium temperature described in Section 2.1.3 and s_{damp} is the damping parameter [days^{-}1] for the zonal mean flow given by Equation (2.10), while s_{damp}^{spec} is the damping parameter for the waves given by

$$s_{damp}^{spec} = \begin{cases} \frac{1}{k_a} + \cos^4(\varphi)\,\frac{1/k_s^{spec} - 1/k_a}{\sigma - \sigma_b} & \text{below } \sigma_b \\ \frac{1}{k_a} & \text{above } \sigma_b \end{cases}$$

where $k_s^{spec} = 0.5$ days, $k_a = 40$ days and φ is latitude, and $\sigma_b = 0.4$ (corresponding to 400 mb) for the run described in Chapter 4.

Chapter 3

Comparison of the Model Data to ERA40 Reanalysis Data

3.1 Introduction

A major part of this thesis makes use of the idealized model introduced in Chapter 2. This chapter examines how the model compares to reanalysis data in order to get a better understanding of the representation and the limitations of stratospheric sudden warmings observed in the model study.

Model studies provide an excellent method to limit the representation to the processes which are believed to be crucial for the studied phenomenon. Stratospheric sudden warmings are observed in purely dynamical models reduced to a dynamical core, i.e. the model does not include radiation physics, chemistry or biology and is purely based on the primitive equations. Simpler 1- and 2-dimensional models have been shown to exhibit phenomena comparable to sudden warmings as described in Chapter 1, however for the present research objectives outlined in the following chapters the higher dimensionality is important. Physical and chemical processes (e.g. ozone chemistry and transport, a major factor determining the mean state of the stratosphere) are not explicitly represented, but partially accounted for in the relaxation to a zonally symmetric equilibrium temperature profile. This representation can be shown to be reasonably realistic since the stratospheric large-scale dynam-

ics comes close to an interplay between a relaxation towards radiative equilibrium disrupted by wintertime wave disturbances forced from the troposphere.

In addition, models provide the possibility for the analysis of cause and effect, since a forcing can not be deliberately placed or removed. The use of models for the study of the impact of specific stratospheric forcings onto the troposphere has in more depth been investigated with this particular model in e.g. Reichler et al. (2005), Ring and Plumb (2007), Ring and Plumb (2008), Chan and Plumb (2009), and Butler et al. (2010).

Since stratospheric sudden warmings are relatively rare events considering the time span that observations exist for (satellite observations are currently available for just over 30 years, with good but more spotty radiosonde and other observations dating about 20 years further back), a large number of model warmings allows for more significant statistics once it is verified that the model warmings indeed exhibit the features desired to be examined in the real atmosphere.

3.2 Reanalysis Data

In order to better understand the above noted benefits as well as the shortcomings of the model in simulating sudden warmings, the analysis of the model warmings is complemented by an analysis of sudden warmings in reanalysis data. Specifically, the Northern Hemisphere sudden warmings observed in ERA40 reanalysis data are compared to the topography run, since this run is the most realistic run performed for this thesis. A more detailed description of the topography run can be found in Chapter 2. The reanalysis dataset used in this chapter is the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalyses ERA40.

The ERA40 reanalysis project (Uppala et al., 2006) includes 23 years of data during the satellite era from 1979 to 2001. Data is provided at 2.5° horizontal resolution and on 23 vertical levels up to 1 hPa, of which 12 levels are located above 200 hPa, yielding an improved stratospheric resolution as e.g. compared to the NCEP reanalysis (Kalnay et al., 1996), which uses 17 vertical levels with a top level at 10 hPa and

of which 7 levels are above 200 hPa.

The ERA40 setup compares to the model predominantly used in this thesis, which (compare to section 2.1) uses a 2.8° horizontal resolution on 40 vertical levels up to 1 hPa, of which 22 levels are located above 200 hPa (excluding the sponge layer which includes 6 layers above 1 hPa).

ERA40 includes 15 major stratospheric sudden warmings during the period used, of which 7 are split warmings and 8 are displacement warmings according to the criterion developed in Charlton and Polvani (2007). This yields a frequency of about one warming every 200 days, depending on the definition of the winter season. This can be compared to the average frequency of warmings computed for the model runs in Table 2.1.

Figure 3-1 shows the zonal mean zonal wind for the ERA40 reanalysis, averaged over all winters (Nov 1 - March 31) 1979-2001. It can be compared to Figures 2-3 to 2-6 in order to get an estimate of the difference of the mean state of the reanalysis data versus the model data. Note that in the reanalysis data, the Northern Hemisphere is the winter hemisphere, while in the model data, the Southern Hemisphere is the winter hemisphere.

In addition to the ERA40 reanalysis, ERAinterim data (Dee et al., 2011) is used for several applications in this thesis. ERAinterim is a continuation and improvement of the ERA40 dataset ranging from 1989 onwards, at a higher resolution and with additional improvements. For the applications in this chapter, however, ERA40 is sufficient.

In addition to there being a much larger sample size for the model warmings as compared to the ERA40 warmings, the ERA40 warmings show a stronger variability among the observed warmings. This is mainly due to the different wave forcing mechanism: In the real atmosphere, Northern Hemisphere planetary-scale Rossby waves are mainly forced by surface topography and zonal surface heating asymmetries, yielding a strong wave-1 as well as wave-2 forcing, while the model run exclusively forces zonal wave-2 at the surface. Wave-2 forcing in the model yields a realistic frequency and strength of sudden warmings as described in Section 2.2.2. This setup yields

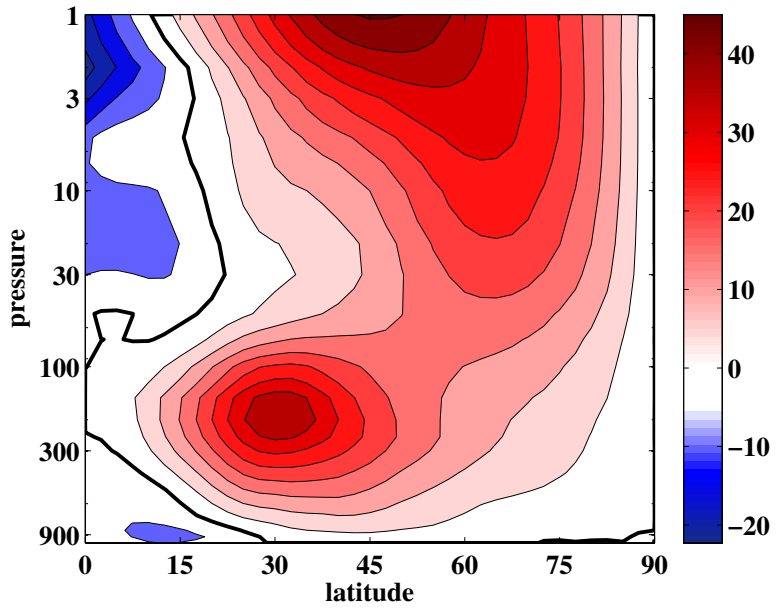

Figure 3-1: ERA40: zonal mean zonal wind in ms^{-1}, the contour interval is $5\,\mathrm{ms}^{-1}$. The zero wind line is printed in bold, westerly winds are shaded in red, easterly winds are shaded in blue. The wind is averaged over all winters (Nov 1 - March 31) 1979-2001. When comparing to the model wind profiles in Figures 2-3 to 2-6, note that the Northern Hemisphere represents the winter hemisphere for the ERA40 data, while for the model runs the Southern Hemisphere is the winter hemisphere.

only split warmings (with a dominant wave-2 pattern at the time of the warming) in the model, while in the real atmosphere both displacement (corresponding to the dominant wave-1 pattern at the time of the warming) and split warmings are about equally present. This thesis will not in detail examine the differences between these two types of warmings, but concentrate on split warmings since those are the ones present in the model study. Other studies have in depth considered the differentiation between warming types [e.g. Charlton and Polvani (2007), Martius et al. (2009), Bancalá et al. (2012)].

3.3 Criteria for the Identification of Sudden Warming Events

Different criteria have been developed in order to determine the timing and magnitude of an observed event, as well as in order to distinguish if a sudden warming happens in the first place. Several criteria will be compared here, as they have been widely used in the literature and since some of them will be employed in the remainder of this thesis. The large number of warmings in the model run will provide a sufficiently large sample to compare the different criteria and to robustly characterize tropospheric effects of sudden warmings.

The best known and most widely used criterion was developed by the World Meteorological Organization (McInturff, 1978), hereafter the *WMO criterion*. According to the WMO criterion, a *major* stratospheric sudden warming has to fulfill the following criteria:

- The zonal mean zonal wind at and poleward of 60° latitude at a height of 10 hPa reverses from westerlies to easterlies.
- The latitudinal temperature gradient at 10 hPa or below and poleward of 60° latitude reverses.

McInturff (1978) defines a minor warming as a significant warming event in the stratosphere over the period of a week or less which does not meet the criteria for a major

Table 3.1: ERA40 sudden warming central dates for 1979 - 2001 in the Northern Hemisphere identified according to the WMO criterion. The warming type is identified according to the algorithm by Charlton and Polvani (2007).

Central date	**Warming type**
22 Feb 1979	splitting
29 Feb 1980	displacement
04 Mar 1981	displacement
04 Dec 1981	displacement
24 Feb 1984	displacement
01 Jan 1985	splitting
23 Jan 1987	displacement
07 Dec 1987	splitting
14 Mar 1988	splitting
21 Feb 1989	splitting
15 Dec 1998	displacement
26 Feb 1999	splitting
20 Mar 2000	displacement
11 Feb 2001	splitting
30 Dec 2001	displacement

warming, implying that minor warmings occur more frequently.

Table 3.1 lists all Northern Hemisphere major sudden warmings (identified according to the WMO criterion) which are observed in ERA40 data for the years 1979 to 2001. The date listed corresponds to the day that the WMO criterion was first fulfilled. The second column lists the dominant wave number responsible for the warming, where a *splitting* corresponds to a warming where wave-2 is dominant at the time of the warming, and a *displacement* corresponds to a warming with a dominant wave-1 signature. The wave number criterion was identified according to the algorithm by Charlton and Polvani (2007). For the period 1979 - 2001 which is considered in this chapter, the WMO criterion identifies 15 sudden warmings, of which 7 are split events and 8 are displacement events.

Using the same algorithm for NCEP reanalysis data (Kalnay et al., 1996) instead yields that several of the listed warmings exhibit a small change in central warming date, a different classification into major and minor warming or into splitting or displacement (see Charlton and Polvani (2007) for a complete list of these cases).

This variability indicates that either the classification by the WMO criterion could be improved, or that sudden warmings cannot uniquely be divided into criteria of magnitude (i.e. *major* or *minor* warming) or longitudinal structure (i.e. *splitting* or *displacement* warming). Coughlin and Gray (2009) confirm that there is no well-defined threshold between minor and major warmings, nor between split or displacement events. However they also find that if a threshold is to be defined, it turns out to be remarkably close to the WMO wind reversal criterion. It is however important to note that sudden warming events are well separated from non-events as well as from cooling events (Polvani and Waugh, 2004).

While the WMO criterion captures the essence of the warming, various other criteria are used throughout the literature, many of them involving the annular mode index time series. Similarly, the criterion used throughout this thesis involves the principal component (PC) time series of the dominant Empirical Orthogonal Function (EOF) mode of the winter hemisphere zonal mean zonal wind at the 10 hPa level, where a warming is defined based on the deviation of the PC time series from its mean. Using only latitudes 20°-90° instead of 0°-90° yields only insignificant changes in the composites. For the ERA40 data, the seasonal cycle was removed before computing the principal component by removing the longterm daily mean wind over all considered years. For the model data, there is no seasonal cycle, so instead the mean of the entire wind time series is removed before computing the principal component. Criteria using -2 or -3 standard deviations of the principal component will be used in this thesis, depending on the effect which is studied with the respective criterion. The stronger criterion using -3 standard deviations (hereafter: PC3) captures events which penetrate further down and which are therefore more likely to have an influence on the troposphere. A warming is defined immediately after the criterion is fulfilled, and a warming is not detected for at least 45 days after the criterion is fulfilled, following Gerber and Polvani (2009). As expected, the events detected with the PC3 criterion yield fewer warmings, but they are all contained within the events detected using the PC2 criterion (using -2 standard deviations), while they are in several cases picked up a couple of days earlier in the PC2 criterion. An additional criterion using

the local minimum of the PC time series over 81 days in cases where the time series falls below 3 standard deviations was also used but it was found that it does not have a strong impact on the composites, as described in Section 5.4.2. While Baldwin and Dunkerton (2001) used the PC3 criterion to identify sudden warmings in their run, Gerber and Polvani (2009) determine that the PC2 criterion best represents a balance between a realistic representation of sudden warmings and reliable statistics in terms of the number of model events for their model run (equivalent to the topography run in this study).

3.4 Analysis of Stratospheric Sudden Warming Composites

3.4.1 Comparison of Sudden Warming Criteria

In order to visualize the difference between the above described criteria, composites of zonal mean zonal wind for sudden warming events were computed for both the ERA40 reanalysis as well as for the topography model run using different criteria: the WMO criterion of wind reversal at 60° latitude and at a height of 10 hPa is compared to the PC2 criterion. It will be shown that the evolution of the wind structure is very similar for these two criteria, with larger differences between the model and the reanalysis data than between the results of using the two criteria.

Figure 3-2 shows the composite of zonal mean zonal wind at 60°N from ERA40 data. The wind data is composited with respect to the WMO criterion, i.e. the central date (day 0) of every warming is determined as the day when the zonal mean zonal wind at 10 hPa and 60°N first reverses to easterlies. Figure 3-3 shows the corresponding composite for the PC2 criterion. Figures 3-4 and 3-5 show the corresponding composites for the model time series.

It can be observed for both the model as well as the reanalysis data that the PC2 criterion yields a smoother weakening of the winds leading up to the warming as compared to the WMO criterion, i.e. the WMO criterion yields a stronger precursor

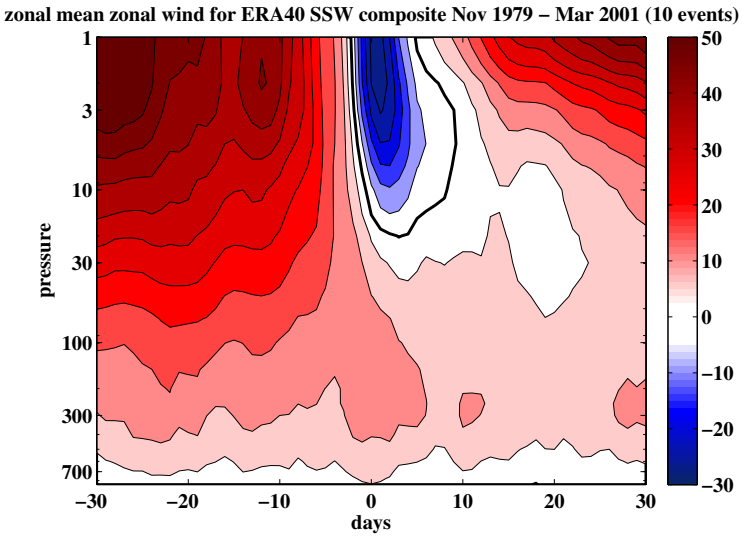

Figure 3-2: ERA40 major warming composite of daily zonal mean zonal wind $\overline{u}$ [ms^{-1}] at 60°N as a function of lag [days] and pressure [hPa] for the sudden warming events between November 30 and March 1 for November 1979 to March 2001. The warmings are identified based on the WMO criterion. This criterion yields 10 events for the considered time period. Day 0 corresponds to the day when the criterion is fulfilled. The bold line corresponds to the zero-wind line ($\overline{u}$=0 ms^{-1}). The contour interval is 5 ms^{-1}.

structure. Since the WMO criterion is a stronger criterion than the PC2 criterion (i.e. the WMO criterion yields less but stronger warmings than the PC2 criterion), this indicates that stronger events need a stronger precursor structure, as they correspond to a stronger weakening of the wind.

Both considered criteria and both the model and the reanalysis show that a precursor structure is necessary for a strong warming to occur. The winds at the vortex location are weakened considerably throughout the month leading up to the warming, and all figures show a deceleration of the winds along with a descent of the maximum wind speeds at 60° latitude. Both the model run and the reanalysis data start out with vortex wind speeds of 45-55 ms^{-1}. The warmings themselves are characterized by a strong enhancement in wind deceleration over the 5 - 10 days before the warming and a slow recovery back to a stronger vortex after the warming event.

The WMO criterion exhibits a stronger response to the warming in the lower stratosphere in both the model and the reanalysis data, characterized by a consistent

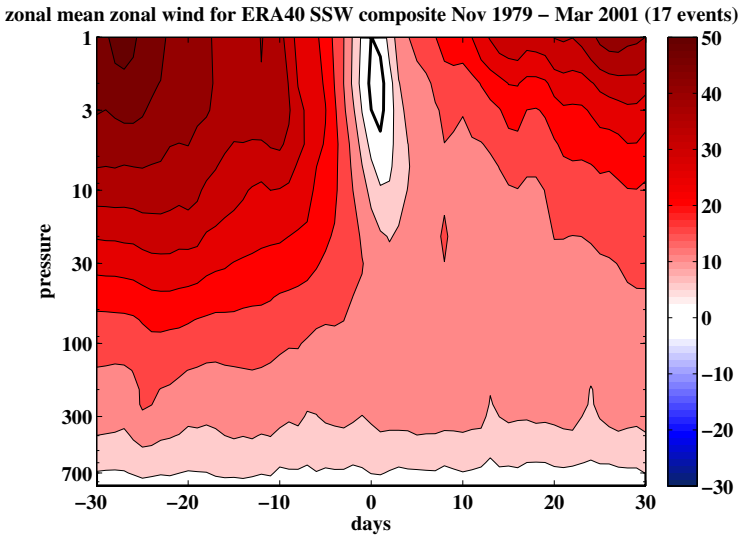

Figure 3-3: Same as Figure 3-2 but using the PC2 criterion, yielding 17 events.

and long-lasting deceleration of the lower stratospheric zonal mean winds right after the warming. This indicates that events with a stronger signature at 10 hPa are able to penetrate further down in the atmosphere.

Considering the actual events that were picked up by the different criteria shows that for the ERA40 data events from the WMO criterion are all represented in the PC2 criterion.

Instead using the PC3 criterion yields 55 events for the topography run (the composite will be shown in Figure 5-2), which can be compared to the 58 events detected using the WMO criterion. Investigating further, 40 of these 55 events detected by the PC3 criterion correspond exactly to events identified by the WMO criterion. This indicates that the PC3 criterion is able to pick up strong and deeply penetrating events, a fact that will be utilized for the analysis in Chapter 5.

Overall, the PC2 and PC3 criteria are shown to be in reasonably good agreement with the WMO criterion, and they pick up realistic representations of stratospheric sudden warmings. Throughout the remainder of the thesis, the PC criteria will be employed to identify sudden warmings for the following reasons:

- The PC criterion is a more holistic criterion as compared to the WMO criterion

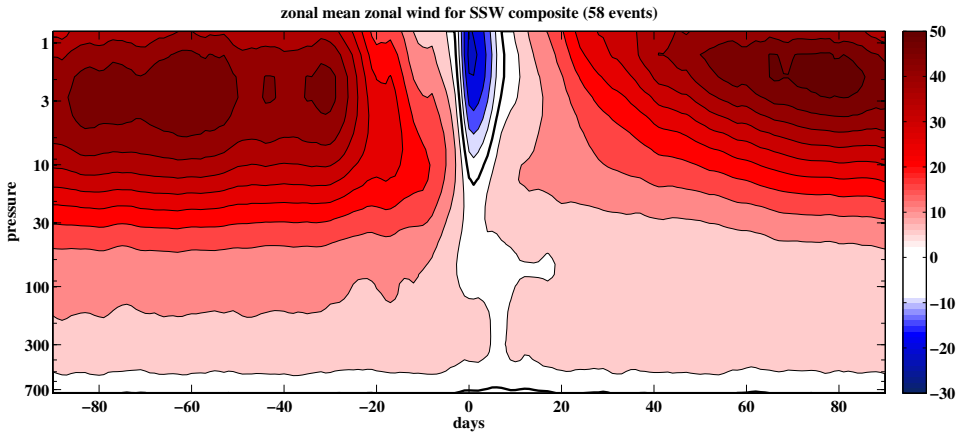

Figure 3-4: Topography run composite time series of daily zonal mean zonal wind at 60°S of all sudden warming events based on the WMO criterion yielding 58 events. Day 0 corresponds to the day when the criterion is fulfilled. The black solid line is the zero-wind line. Contour interval is $5\,\mathrm{ms}^{-1}$. The composites were chosen to extend further than for Figures 3-2 and 3-3 since there is no seasonal cycle in the model composite which can alter the flow during the period leading up to the warming or during the recovery period.

by encompassing the variability of the entire hemisphere as opposed to just a small part of the winter stratosphere. It is in addition consistent with how much of the literature identifies sudden warmings, which makes the present study easier to compare to other studies.

- The PC criterion can be relaxed and strengthened in a much more objective way as compared to the WMO criterion by simply increasing or decreasing the standard deviation at which events are picked up.

- It has consistently been shown (compare e.g. Chapter 5) that the annular mode index is strongly coherent throughout the stratosphere. In addition, the stratospheric annular mode is closely connected to the tropospheric annular mode, and the co-variability between these modes is especially strong during winters with a strongly disturbed stratosphere, i.e. during sudden warming activity.

In general, the warming composites are comparable for the reanalysis and the topography model data, while the different criteria for identifying sudden warmings are

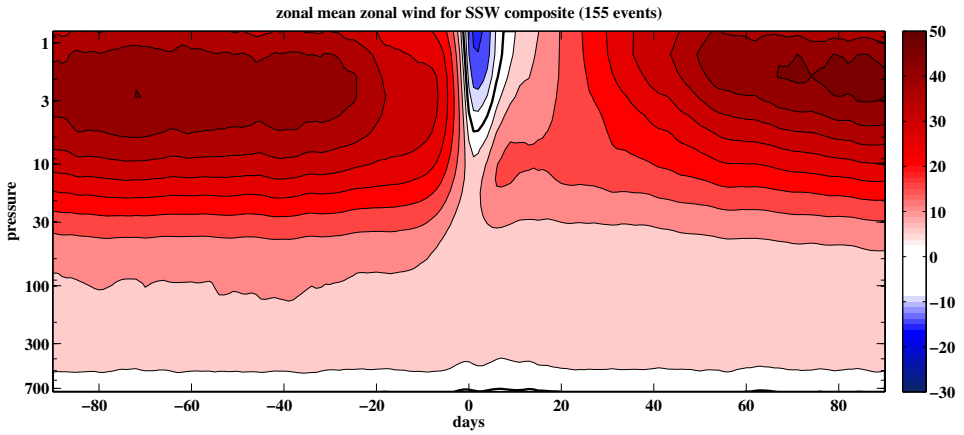

Figure 3-5: Same as Figure 3-4 but based on the PC2 criterion yielding 155 events.

consistent at picking up the warmings with only minor differences between the criteria. Note that the composite of the ERA40 events consists of much fewer warmings as compared to the model data, thereby smoothing out the precursor structure as well as the warmings themselves in the model composite. The PC criteria are straight forward to relax and strengthen and will be used in the remainder of this thesis.

3.4.2 Comparison of Eliassen-Palm Flux Composites

Using the knowledge that stratospheric sudden warmings are forced by wave flux from the troposphere, the wave fluxes into the stratosphere are here compared in order to get a better understanding of the difference in wave forcing during sudden warmings between reanalysis and model data. All composites were computed according to the PC2 criterion. The vertical EP flux is used as a proxy for vertical wave propagation, given by

$$F_z \quad = \quad \rho_0 f a \cos\varphi \frac{\overline{v'\theta'}}{d\bar{\theta}/dz}\,, \tag{3.1}$$

where $\rho_0(p)$ is density as a function of pressure, φ is latitude and z is height, f is the Coriolis parameter, a is Earth's radius, $\overline{v'\theta'}$ is the meridional heat flux, and θ is

potential temperature.

Figure 3-6 shows the winter average of the vertical EP flux (hereafter: F_z) scaled by density divided into the contributions by zonal wave numbers 1, 2, and 3, and averaged over all winters between November 1979 and March 2001. Wave-1 exhibits the strongest F_z into the stratosphere, while F_z for wave-3 is smaller by about one order of magnitude. Only the tropics and part of the polar latitudes show downward F_z in the winter average.

Figures 3-7 and 3-8 show the anomalous F_z for sudden warming composites of ERA40 data. The anomalies are computed with respect to the seasonal cycle averaged over the years 1979 - 2001 before compositing the sudden warming events. In the stratosphere, wave-1 dominates F_z in the lead-up to the warming, while wave-2 grows dominantly around the time of 10 to 5 days before the warming (Figure 3-7). The stratospheric winds are not weakened continuously by the EP fluxes, but by precursors in terms of several strong wave bursts before the occurrence of a major warming. Often, a minor warming or other strong wave events are observed before a major warming, and these precursors have been consistently observed in the literature: Labitzke (1981) has found strong wave-1 precursors about 2 to 3 weeks before the warming for many of the observed warmings in the Northern Hemisphere. Wave-1 is more likely to be able to propagate into the strong vortex winds to yield an initial weakening due to the criterion by Charney and Drazin (1961), while wave-2 often picks up at a later stage of the weakening, especially during split events. Labitzke (1981) also found that wave-1 and wave-2 tend to be anti-correlated in their growth, however it is not resolved if this is simply a matter of polar vortex geometry (i.e. the vortex assuming a shape which in wave number space is identified as a more distinctly wave-1 or wave-2 pattern) or if the waves are not able to amplify simultaneously [see also Schoeberl (1978)]. After the warming occurs, a significant decrease in F_z flux can be observed, consistent with both observational data [e.g. Limpasuvan et al. (2004)] and modeling data [e.g. Gerber and Polvani (2009), Polvani and Waugh (2004)].

In the troposphere, F_z looks more messy, as expected. Again, wave-1 dominates in terms of upward flux before the warming, while wave-2 strongly contributes to the

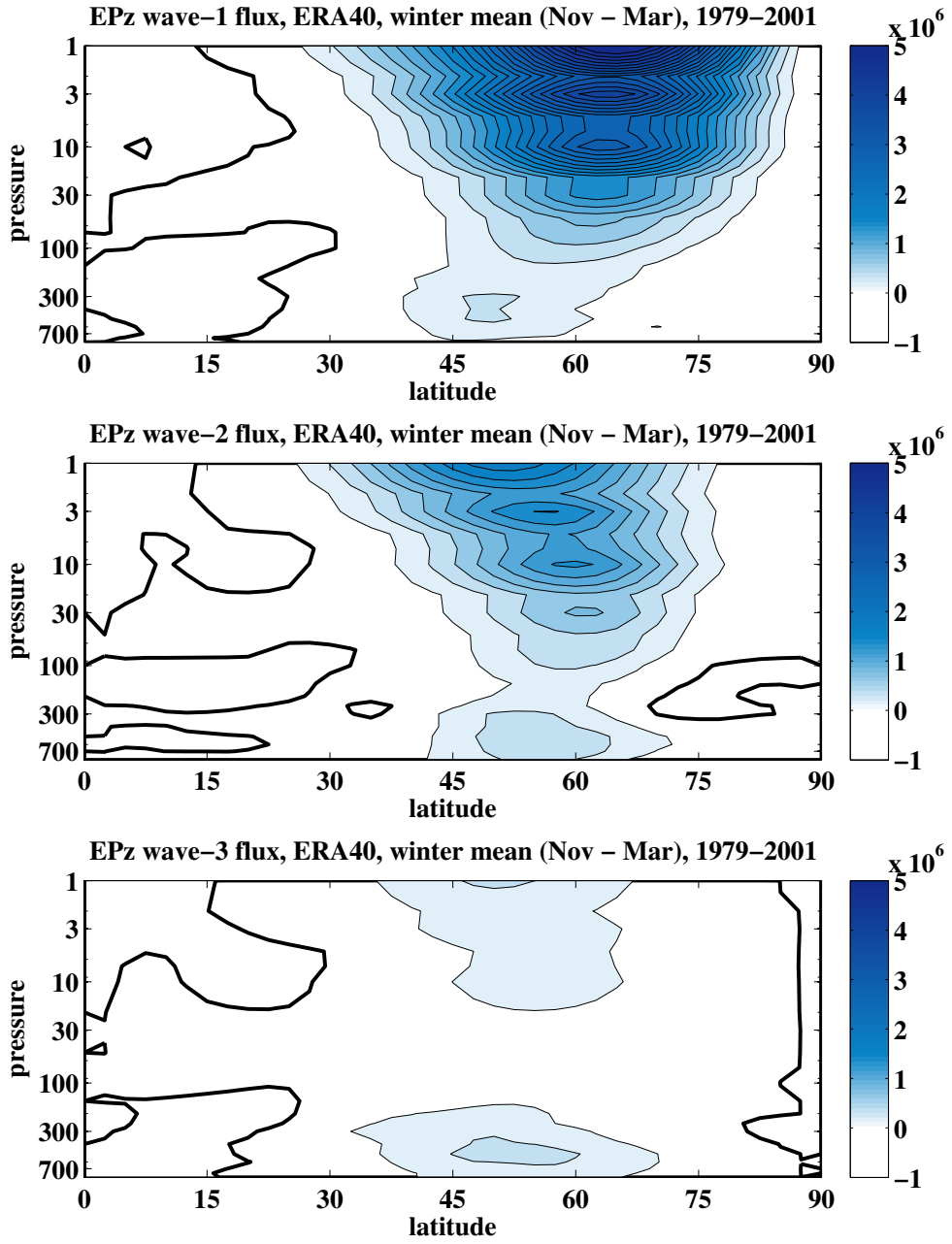

Figure 3-6: Vertical component of the EP flux (from Equation (3.1)) weighted by $\frac{1}{\rho_0}$. The flux is averaged over the winter season (Nov 1 - Mar 31) between Nov 1979 - Mar 2001. The panels show F_z divided into different wave numbers 1 (top), wave-2 (middle) and wave-3 (bottom). Units: $\mathrm{m^3 s^{-2}}$. Contour interval: $2{\cdot}10^5$ $\mathrm{m^3 s^{-2}}$.

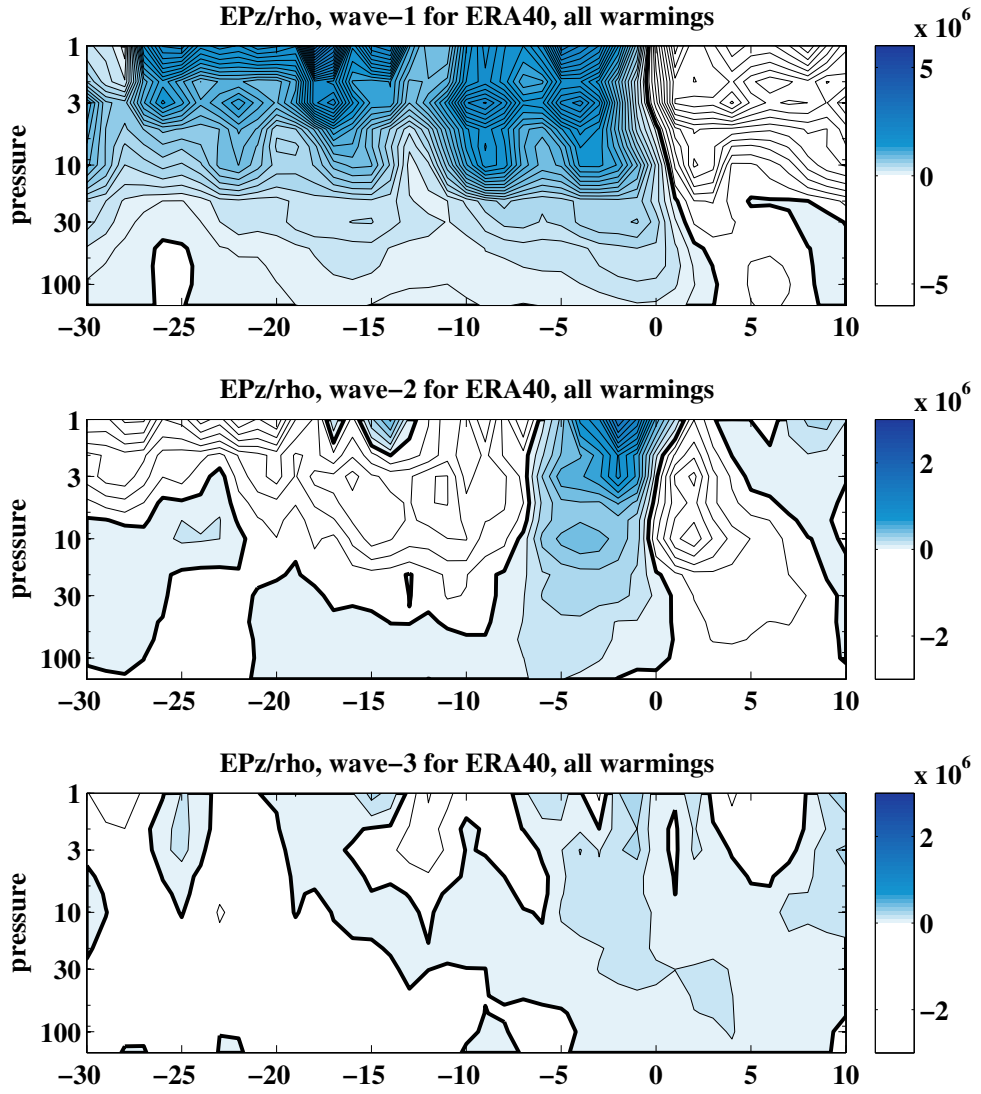

Figure 3-7: Composite of the anomalies with respect to the seasonal cycle of the vertical component of the EP flux (F_z) weighted by $\frac{1}{\rho_0}$ for the stratosphere only and composited for all sudden warmings observed according to the PC2 criterion as a function of lag [days] around the central date of the sudden warming composite. The EP flux is divided into wave numbers 1 (top), wave-2 (middle) and wave-3 (bottom) and integrated over 40°-80°N. Units: $\mathrm{m^3 s^{-2}}$. Contour interval: $10^5\,\mathrm{m^3 s^{-2}}$.

wave flux right before the warming. The weakening of the lower stratospheric winds as observed in Figure 3-3 may allow wave-3 to propagate into the otherwise strong winds, yielding a weak wave-3 amplification before the warming in the stratosphere.

So far no distinction has been made between warmings dominantly induced by wave-1 or wave-2, i.e. displacement or split warmings, and both wave numbers are present in the composites. Wave-1 is dominant throughout the time leading up to the warming, while wave-2 tends to grow later on, probably due to the weakened winds allowing for propagation of smaller wave numbers. For comparison to the model it is interesting to take a look at split warmings exclusively, since all warmings observed in the model are split warmings.

Figures 3-9 and 3-10 show the same composites as 3-7 and 3-8 using the PC2 criterion for ERA40 reanalysis, but only including split warmings, i.e. 6 events for the period between November 30 and March 1 for November 1979 to March 2001 (compare to Table 3.1). An increase in the maximum wave flux can be observed for the split events as compared to the composites including both split and displacement events. The wave-1 precursors around day -25 and day -15 have strengthened with respect to the composite including both displacements and splits, while the wave-2 flux preceding the warming by a couple of days has strengthened as well. While one would expect the strengthening of the wave-2 flux right before the warming from the definition of the splitting events, the result for the precursor structure is more surprising. Split warmings exhibit stronger wave-1 precursors than the figures displaying the average over both split and displacement warmings. This may indicate that wave-2 is dependent on a strong weakening of the wind by wave-1 (which is able to propagate into stronger winds than wave-2) before wave-2 is able to propagate. However it has to be kept in mind that this composite includes only 6 split events and stronger statistics would be required to determine the significance of this finding. However it is known from e.g. Labitzke (1981) that wave-1 is dominantly present in the real world and that it often grows before sudden warmings, also as a precursor for split warmings. On the other hand, wave-2 often does not grow much before displacement warmings, as can be seen by comparing Figures 3-7 and 3-9.

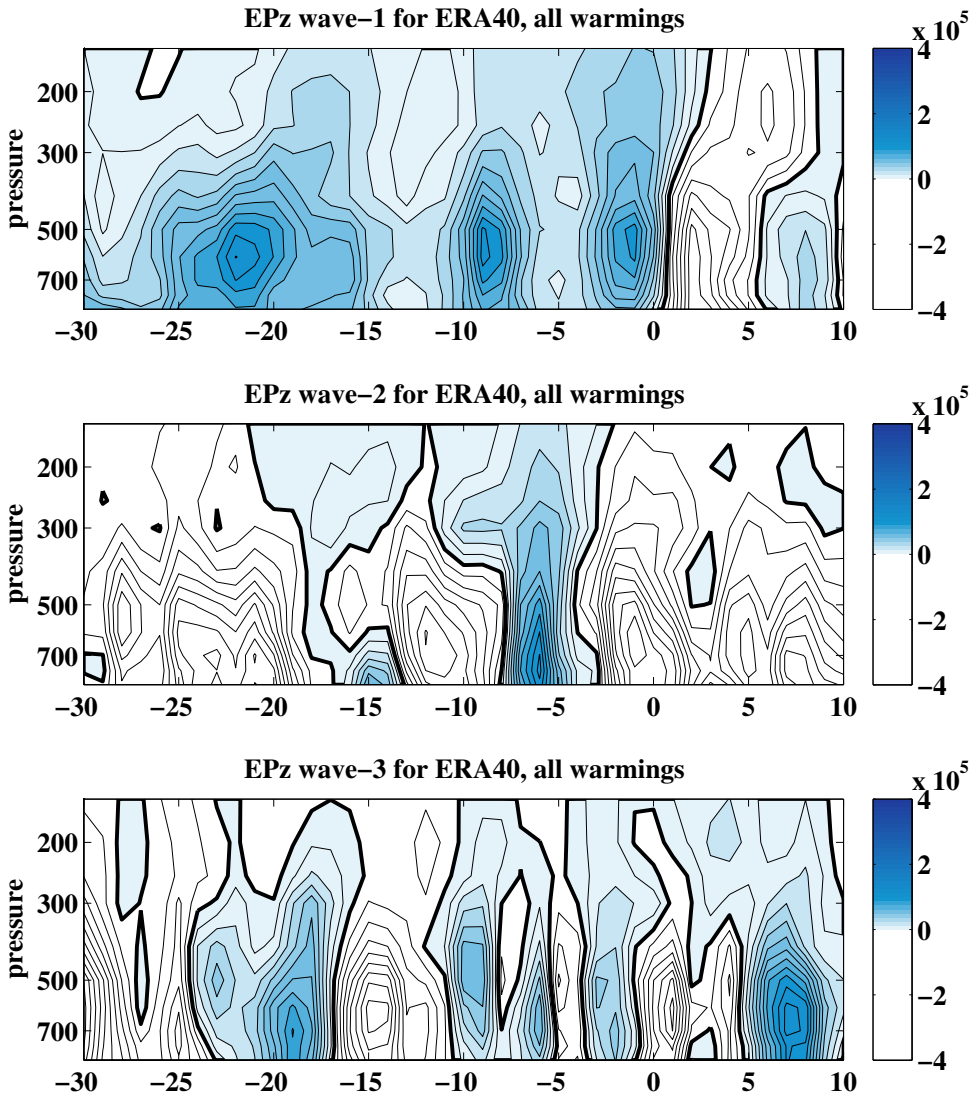

Figure 3-8: Same as Figure 3-7 but for the troposphere only and unweighted, i.e. as given by Equation (3.1). Units: $\mathrm{kg\,s^{-2}}$. Contour interval: $10^4\,\mathrm{kg\,s^{-2}}$.

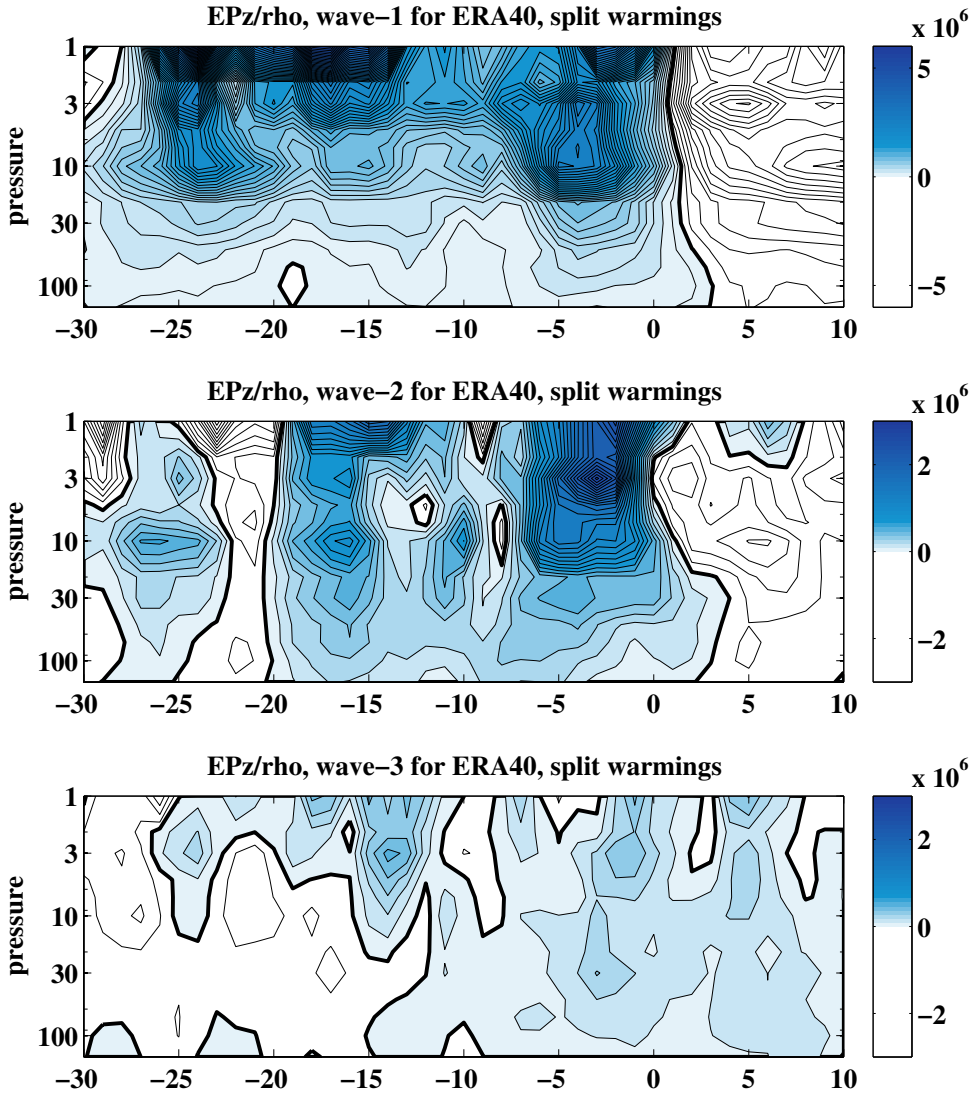

Figure 3-9: Same as Figure 3-7 but for split warmings only. Units: $m^3 s^{-2}$. Contour interval: $10^5\, m^3 s^{-2}$.

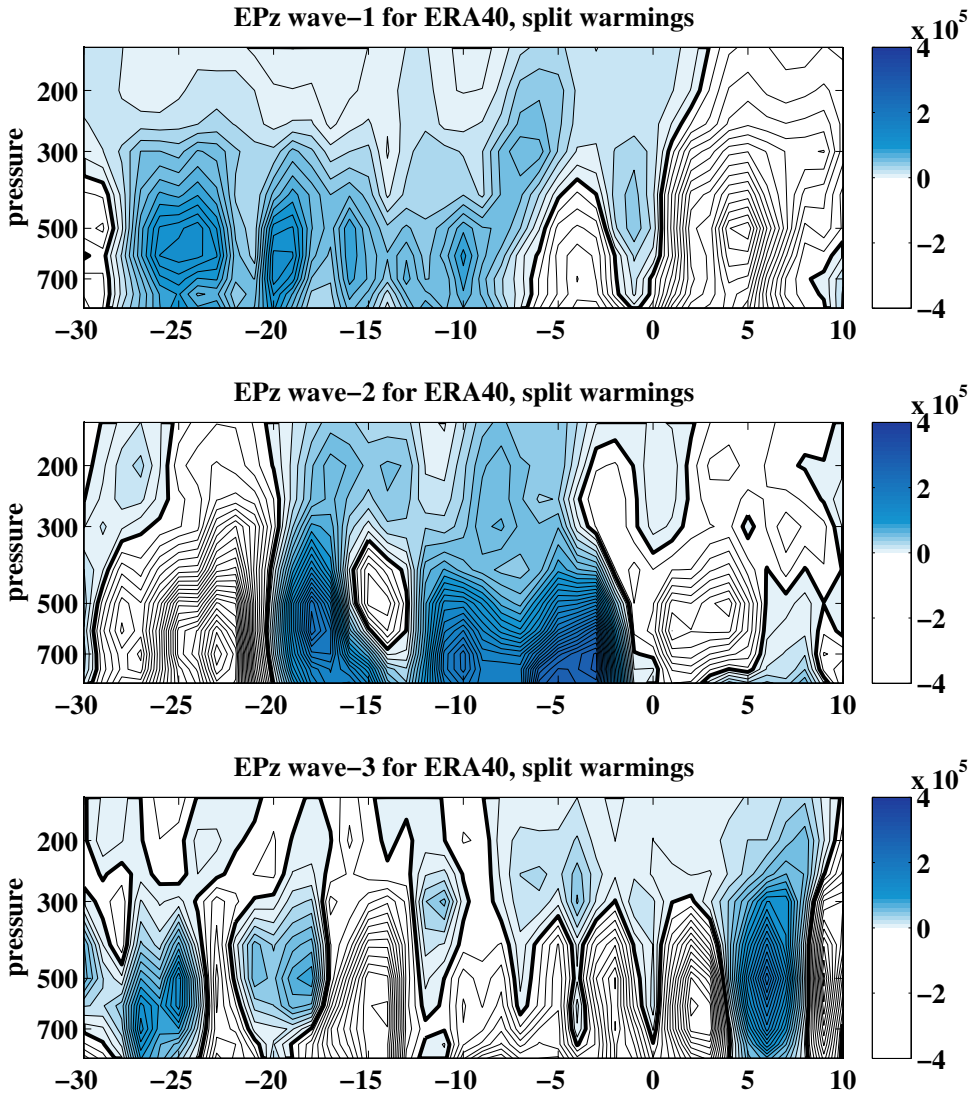

Figure 3-10: Same as Figure 3-8 but for split warmings only. Units: $\mathrm{kg\,s^{-2}}$. Contour interval: $10^4\,\mathrm{kg\,s^{-2}}$.

The sudden warming composites from reanalysis data are now compared to the model warming composites. Note that these figures are plotted for the Southern Hemisphere, which represents the winter hemisphere. Figure 3-11 shows the vertical EP flux weighted by density for the model run, and averaged over the entire model run. For the model, the strongest upward EP fluxes can be observed for wave-2, since wave-2 is explicitly forced at the model lower boundary, while all other wave numbers are not explicitly forced. Again, F_z fluxes for wave-3 are about an order of magnitude smaller than the dominant wave fluxes. Wave-1 reaches magnitudes about half of wave-2 despite not being forced, which probably can be explained by wave-1 being able to propagate through very strong winds as opposed to larger wave numbers or by local nonlinearities from wave-wave interaction induced by the wave-2 patterns yielding asymmetries of the vortex about the pole.

Figures 3-12 and 3-13 show the anomalous F_z flux for sudden warming composites, again divided into contributions from zonal wave-1, wave-2, and wave-3. The anomalies are computed with respect to the longterm mean of the model run, since there is no seasonal cycle. A very clear dominance of wave-2 can be observed, with the warmings exhibiting both a wave-2 precursor structure as well as a very dominant amplification of wave-2 before the warming. This amplification corresponds to more than a 50% increase in F_z flux for wave-2 as compared to the climatology. Wave-3 contributes very little with wave fluxes an order of magnitude smaller than wave-2, while wave-1 even exhibits negative F_z around the time of the warming.

Wave-2 exhibits large wave fluxes already in the upper troposphere, where strong waves seem to be forced and are able to propagate upward. This can be compared to the positive wave-2 F_z flux in the reanalysis data about 5 days before the warming (Figure 3-8). Wave-3 again shows little contribution to the total wave flux. The collapse of the wave flux after the warming can clearly be observed, although it is slower than in the reanalysis data, where the wave flux collapses a couple of days earlier after the warming.

From both the EP flux time mean figures as well as the sudden warming composites it is clear that wave-3 does not significantly contribute to stratospheric variability in

the topography model run, with wave fluxes generally an order of magnitude smaller than for wave-1 and wave-2 in the upper stratosphere. Wave-2 is considerably stronger as compared to wave-1 in the topography run as it is forced in the troposphere. It contributes to more variability in the stratosphere than wave-1, while however wave-1 peaks further up in the stratosphere or even the mesosphere in the climatological mean. Wave-2 starts decaying around a height of 3 hPa, while wave-1 is able to propagate further up, since the stratospheric winds act as a low-pass filter on wave numbers. The dominant pathways of wave propagation can also be observed in these figures: wave-2 propagates into the tropospheric jet and then picks up again in the stratosphere.

For completeness and for comparison in the later chapters, the same composites are computed for the run truncated to wave-2 as described in Section 2.2.3. The top panel in Figure 3-14 shows the vertical EP flux averaged over the entire run. A comparison to Figure 3-11 (middle panel) shows that the vertical EP flux of wave-2 has strengthened considerably in the troposphere, while it is also more spread out in the troposphere. This difference between the topography run and the truncated run will be further discussed in Chapter 4. The middle and bottom panels in Figure 3-14 show F_z for the stratosphere and the troposphere, respectively, composited over the sudden warmings according to the PC2 criterion, yielding 106 events. Only wave-2 is shown since the other wave numbers are truncated in this model run. The characteristic precursor structure can be observed in the stratosphere, while an amplification of the wave can be observed throughout the model atmosphere several days before the warming, comparable to the topography run as well as the reanalysis data.

Polvani and Waugh (2004) showed that the important or even the determining quantity for the occurrence of a sudden warming is the time integrated heat flux entering the stratosphere. It is therefore not surprising to see the strong precursor structures observed in both the reanalysis data as well as the model data: A strong wave flux right before the warming by itself may not be able to cause a strong warming, but it accounts for the final weakening of the vortex. Instead, a strong vortex such as the one forced in the model run has to be weakened by strong upward EP

fluxes leading up to the warming. In the model, the precursors are dominated by wave-2, while in the reanalysis data, dominantly wave-1 is responsible for precursors.

3.5 Conclusion

While the model warmings are not an exact representation of the real world, it can nevertheless be concluded that the employed relaxation of stratospheric temperatures to a zonal mean equilibrium temperature profile is sufficient to reproduce the relevant dynamics of sudden warmings. The main difference between the model run and the reanalysis data is the different forcing structure, with only wave-2 being forced at the model surface. Wave-2 is also responsible for all warmings observed in the model run. The Northern Hemisphere, however, exhibits a strong surface forcing of both wave-1 and wave-2, as can be observed in the reanalysis data in the results for vertical EP flux. In addition, the climatological polar vortex is located slightly higher up in the reanalysis (compare Figures 2-4 and 3-1), leading to a difference in propagation characteristics of the waves in the upper stratosphere. However, this does not affect the characteristics of the warmings. The traditionally used WMO criterion identifies strong and deeply penetrating events, similar to the PC3 criterion, while the PC2 criterion identifies weaker warmings which on average are not able to strongly influence the tropospheric flow in both the reanalysis and the model run. The weakening of the mean winds happens over a similar time interval and at a similar strength, and both the model as well as the reanalysis exhibit a strong precursor structure. Due to the model setup, these precursors are dominated by wave-2 in the model run, while wave-1 dominates the precursors in the reanalysis, with wave-2 responsible for a smaller part of the weakening of the mean flow. Before the warming occurs, the strengthening of the EP fluxes can be traced down into the troposphere for both the reanalysis and the model run.

After having identified the key limitations and benefits of the model in simulating sudden warmings, the following chapters will analyze different properties of the observed warmings in both the topography run as well as modified model runs.

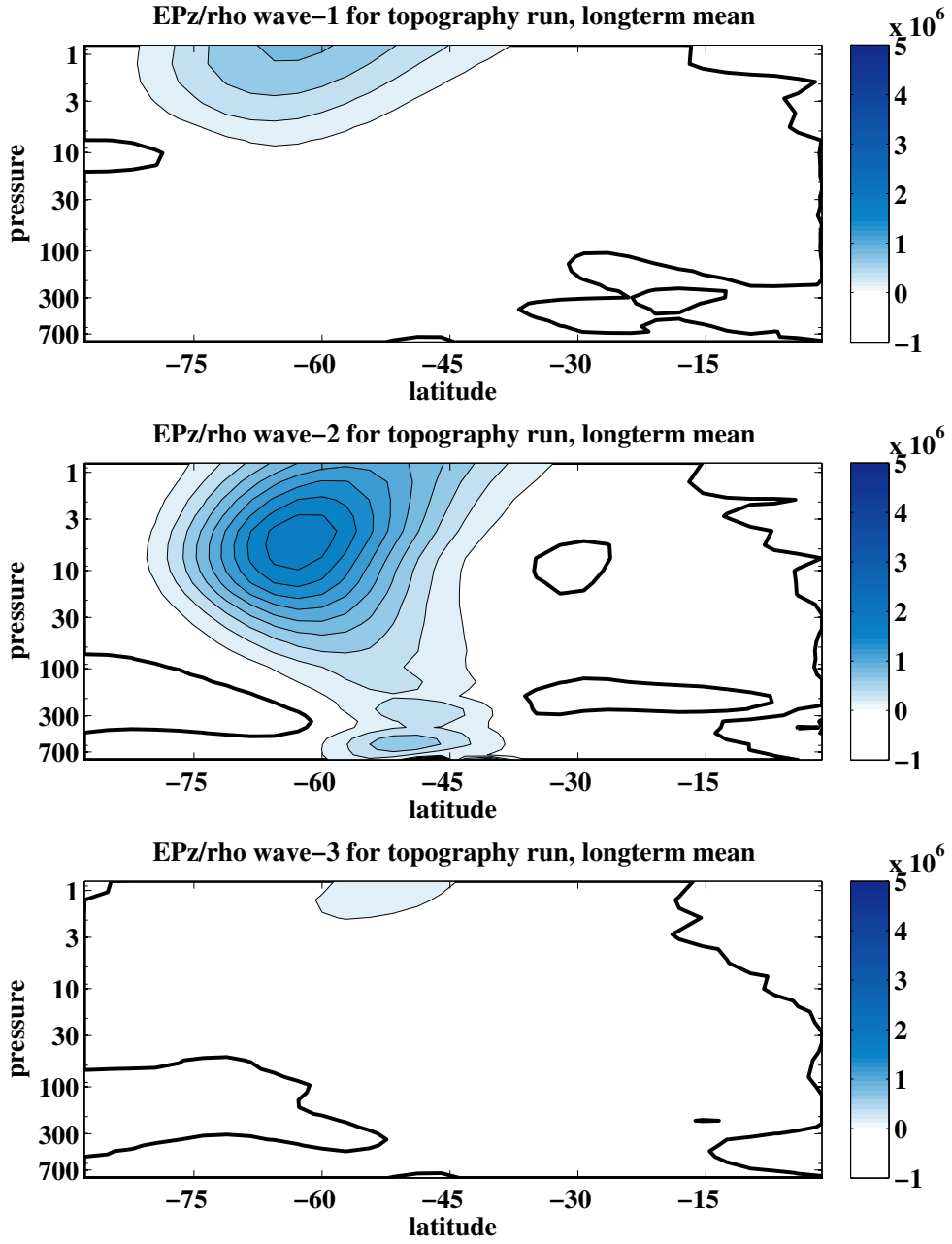

Figure 3-11: Vertical component of the EP flux (F_z from Equation (3.1)), weighted by ρ_0^{-1}. The flux is integrated over 40°-80°S and averaged over the entire model run. The panels show F_z divided into zonal wave numbers 1 (top), wave-2 (middle) and wave-3 (bottom). Units: $\mathrm{m^3 s^{-2}}$. Contour interval: $2 \cdot 10^5\,\mathrm{m^3 s^{-2}}$.

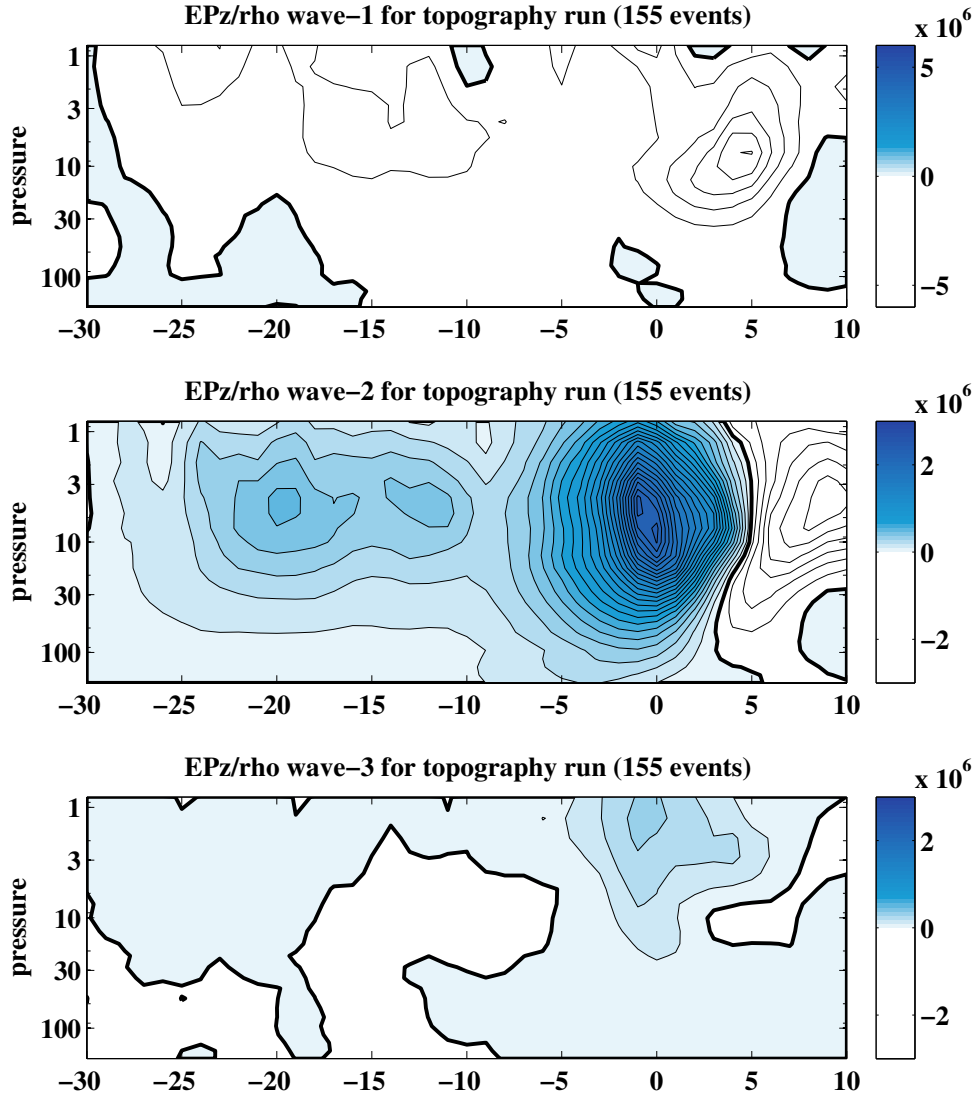

Figure 3-12: Composite of the anomalous vertical component of the EP flux F_z weighted by ρ_0^{-1} for the stratosphere only as a function of lag [days] around the central date of the sudden warming. The fluxes are composited for all sudden warmings observed in the run according to the PC2 criterion. The F_z flux is divided into wave numbers 1 (top), wave-2 (middle) and wave-3 (bottom) and integrated over 40°-80°S. Units: $\mathrm{m^3 s^{-2}}$. Contour interval: $10^5\,\mathrm{m^3 s^{-2}}$.

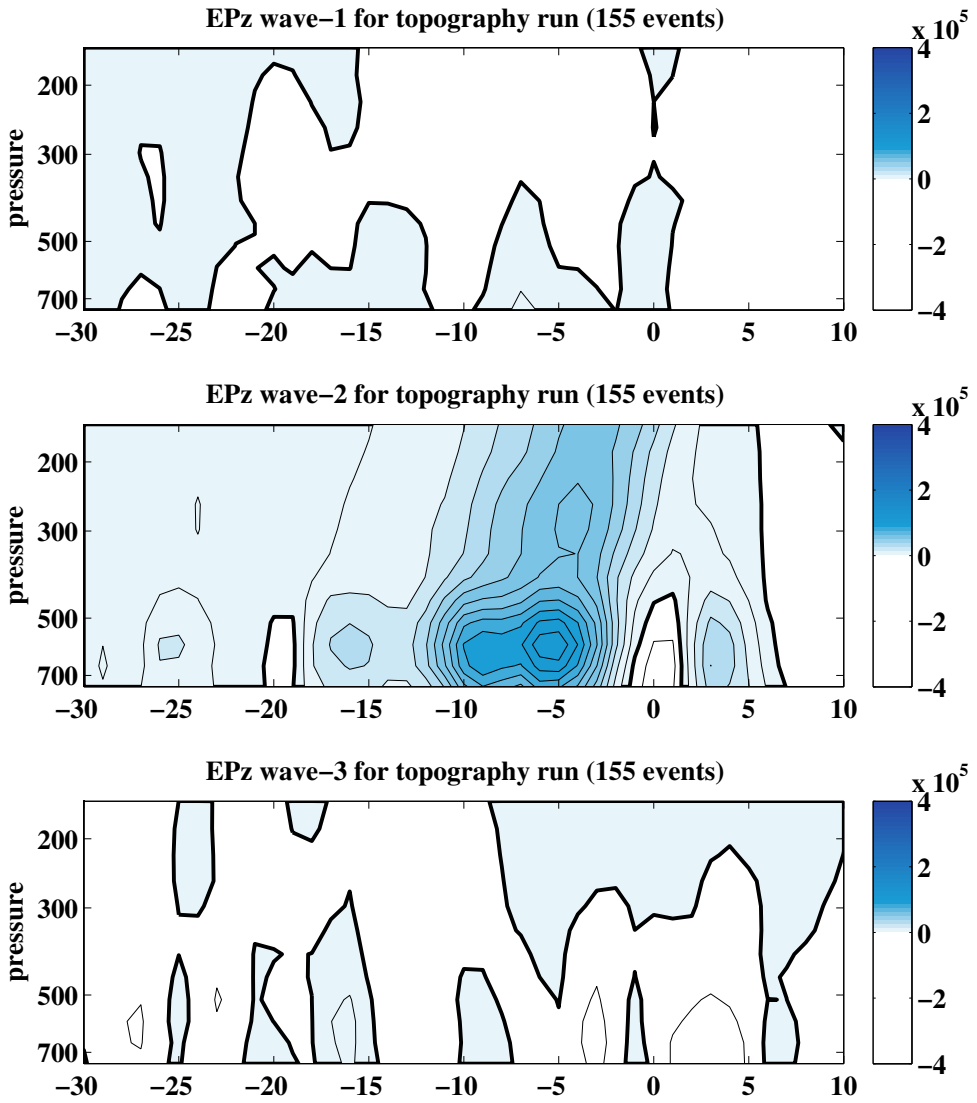

Figure 3-13: Same as Figure 3-12 but for the troposphere only and unweighted, i.e. given by Equation (3.1). Units: $\mathrm{kg\,s^{-2}}$. Contour interval: $10^4\,\mathrm{kg\,s^{-2}}$.

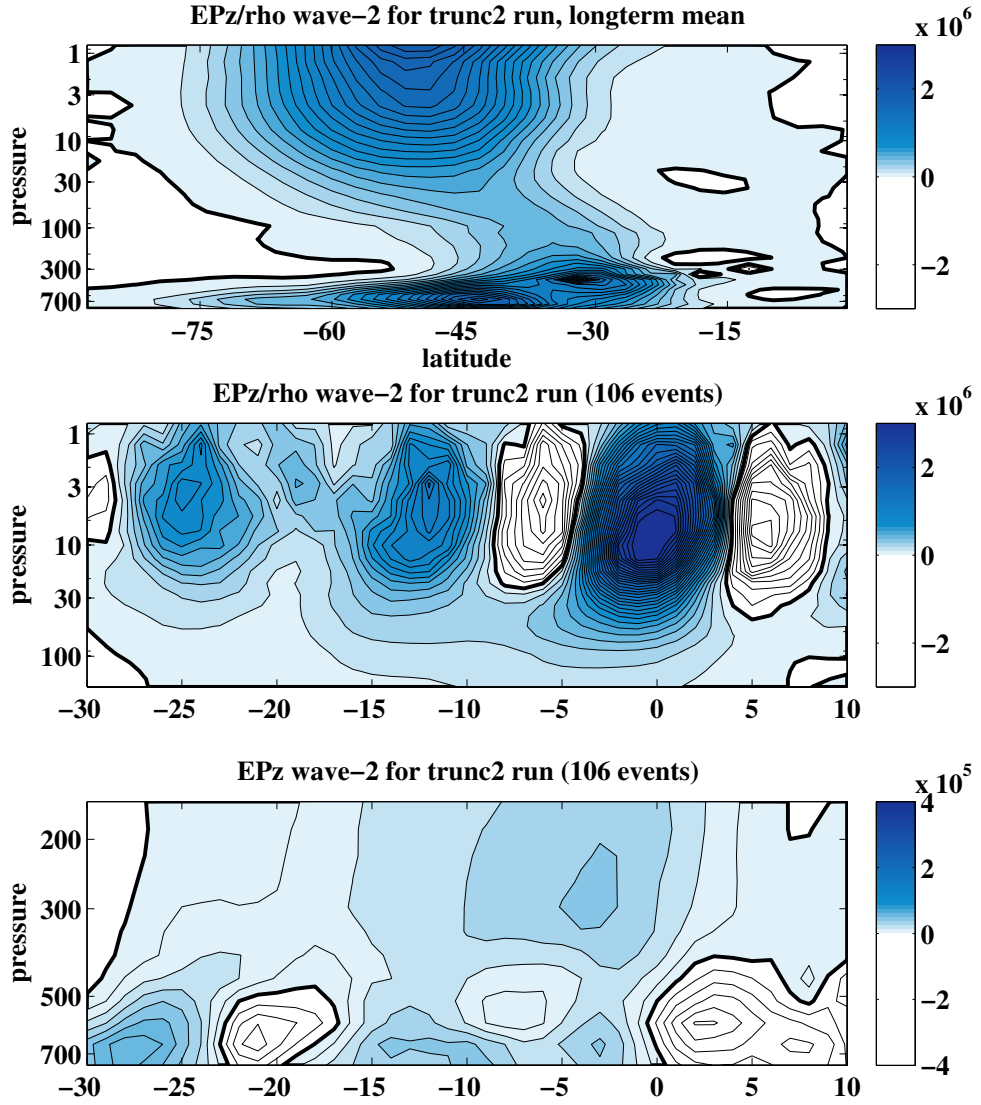

Figure 3-14: Vertical EP flux for the run truncated to wave number 2: Top panel: $\rho_0^{-1} F_z$, compare to Figure 3-11 (middle panel). Contour interval: $10^5\,\mathrm{m^3 s^{-2}}$. Middle panel: $\rho_0^{-1} F_z$ for the sudden warming composite for the stratosphere. Compare to Figure 3-12 (middle panel). Contour interval: $10^5\,\mathrm{m^3 s^{-2}}$. Bottom panel: F_z in the troposphere for the sudden warming composite, compare to Figure 3-13 (middle panel). Contour interval: $10^4\,\mathrm{kg\,s^{-2}}$. The heavy line is located at zero for all panels.

Chapter 4

Traveling Planetary-Scale Waves in the Stratosphere: Tropospheric Causes and Stratospheric Internal Variability

4.1 Introduction

4.1.1 Southern Hemisphere Stratospheric Winter Variability

Quasi-stationary planetary-scale Rossby waves are the dominant driver of sudden warming events and more generally extratropical stratospheric variability in the Northern Hemisphere stratosphere. These waves are caused by longitudinally varying surface structures such as topography and differential heating patterns, as observed in the Northern Hemisphere in the distribution of land and ocean, which yields variations in surface pressure in terms of both changes in height structure and heating. Due to their stationary or quasi-stationary nature, these surface forcings lead to the generation of stationary or quasi-stationary Rossby waves in the Northern Hemisphere. In the Southern Hemisphere, however, strong topography or longitudinally varying heating patterns are largely absent in the midlatitudes due to the zonally symmetric

nature of the Southern Ocean. Only a weak quasi-stationary feature due to Antarctic topography can be observed. The much weaker surface forcing in the Southern Hemisphere is generally believed to be the main reason for the much more quiescent winter stratosphere in the Southern Hemisphere as compared to the Northern Hemisphere.

The Southern Hemisphere stratosphere is dominated by traveling planetary-scale waves with a smaller impact on the zonal mean flow as compared to the stationary waves in the Northern Hemisphere stratosphere. The origin of these traveling waves is not fully understood, but upper and mid-tropospheric mechanisms are believed to contribute to the generation of traveling Rossby waves by e.g. wave-wave interaction of synoptic-scale waves (Scinocca and Haynes, 1998) and tropospheric baroclinic instability of long waves (Hartmann, 1979). Since there is no reason for these mechanisms to produce stationary or quasi-stationary waves (unlike for stationary surface forcings), these processes result in longitudinally traveling planetary-scale waves. While of large amplitude and vertically propagating, the traveling waves are not generally believed to be able to induce sudden warmings, while they likely play a role in the interaction with stationary waves [Plumb (1981), Plumb (2010)], thereby contributing to the forcing of sudden warmings [e.g. Esler et al. (2006)]. The only evidence for a major sudden warming that happened in the presence of only weak zonally asymmetric surface forcing is the Southern Hemisphere sudden warming in September 2002 (the WMO criterion was fulfilled on Sept 26). Figure 4-1 shows a comparison between the zonal mean zonal winds for the Southern Hemisphere averaged over the winters 1989-2010 excluding 2002 (solid black line) as compared to winter 2002 (red line).

This chapter will compare the respective roles of the tropospheric generation mechanisms of stratospheric traveling waves and examine the impact of traveling waves on the stratospheric flow. The hypothetical case of a troposphere with no synoptic eddies will be studied and compared to a troposphere including the full model wave number space in order to explore two extreme cases of the parameter space. While the model run including synoptic eddies is more realistic with respect to Southern Hemisphere winter, it is illuminating to assess the role of the tropospheric mean state as well as the wave generation mechanism in producing stratospheric traveling waves.

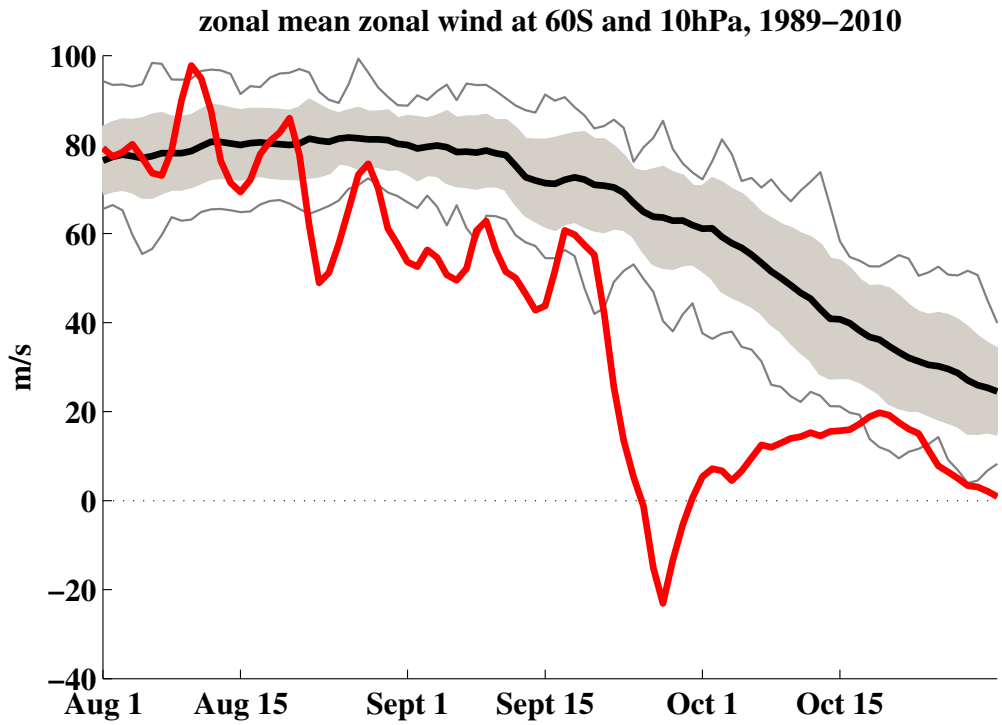

Figure 4-1: Time series of daily zonal mean zonal wind at 60°S and 10 hPa from ERAinterim data. The solid black line shows the mean value for the wind averaged over the winters 1989-2010, excluding 2002. The gray shading denotes the standard deviation, and the thin gray lines denote the daily maximum and minimum values of zonal mean zonal wind for these years. The solid red line depicts daily zonal mean zonal wind at 60°S and 10 hPa for Aug - Oct 2002. Southern Hemisphere winter 2002 deviates strongly from other winters as early as mid-August, and the warming itself is stronger in terms of wind deceleration as compared to typical Northern Hemisphere sudden warmings.

4.1.2 Tropospheric Generation Mechanisms for Stratospheric Traveling Waves

Several mechanisms have been proposed to be able to force traveling waves in the absence of a surface wave forcing: As a first mechanism, Hartmann (1979) suggested the generation of traveling waves by Charney-type baroclinic instability to planetary-scale wave disturbances in the troposphere. Hartmann demonstrated that realistic mean flow states may exhibit instabilities to zonal wave numbers 2 and 3, and that these waves have growth rates that allow them to propagate into the stratosphere. In order for these instabilities to occur, the mean wind has to exhibit a strong vertical shear, and the wave has to be confined to high latitudes along with a comparatively small meridional scale of the wave. Hartmann derived these factors for instability from considering wave growth on a β-plane: An increase in vertical shear enhances relative vorticity advection with respect to the β-term. The magnitude of β is further decreased by moving wave growth to higher latitudes, which will in addition decrease the zonal wavelength of the wave through the decrease in the length of the latitude circle with increasing latitude. In addition, a small meridional scale will enhance the ability for the wave to grow. With a smaller meridional and zonal scale, the wave starts resembling a synoptic-scale wave, making it more likely for the flow to become unstable to these waves.

On the other hand, in order for the wave to be able to propagate into the stratosphere, it needs to extend its horizontal scale as it propagates into the stratosphere in order to fulfill the wave propagation criterion by Charney and Drazin (1961), which shows that vertical wave propagation is dependent on the sum of the squares of the horizontal wave numbers. For Hartmann's experiments in a quasi-geostrophic β-plane model, the critical velocity $u_c = \frac{\beta}{k^2+l^2}$ is the maximum wind speed that a wave of horizontal wave numbers k and l can propagate into at a given latitude corresponding to a given value of β. Hartmann showed that zonal wave-2 exhibits large growth rates as well as deep propagation into the stratosphere, with the meridional scale of the wave in the stratosphere being almost double of the meridional scale in the troposphere.

The strongest limitations for this theory to be relevant for the real world are the above mentioned criteria required for instability, especially the criterion for reaching a sufficient vertical wind shear, as will be shown in the model studies in this chapter.

Scinocca and Haynes (1998) suggested another mechanism for the generation of traveling planetary-scale waves: In numerical experiments they showed that planetary-scale waves can arise from nonlinear wave-wave interactions between synoptic-scale baroclinic eddies in the troposphere, which organize into larger wave packets of planetary scale. Comparable to the model used in this thesis, their model is a primitive equation model with a relaxation to a zonally symmetric equilibrium temperature profile representing a winter stratosphere, with no zonally asymmetric forcing. They confirm the generation mechanism in their model by inserting the eddy fields from the nonlinear model into a linear model. Although they observe strong organization of the waves into planetary wave-2 packets, they find that the waves in the nonlinear model are not able to readily propagate into the stratosphere, explaining about half of the observed wave amplitude in the stratosphere. This discrepancy could arise for several reasons, i.e. the mechanism by itself may not be responsible for the entire wave flux into the stratosphere, and the remaining waves are produced by a different mechanism, or the lower stratospheric winds may not be favorable for wave propagation and may act as a valve for the waves as shown in Chen and Robinson (1992).

4.2 Model Setup

In order to investigate and compare the described mechanisms, we compare several model runs with no surface forcing in order to avoid the forcing of quasi-stationary waves. We partly control the mechanism in the experiments with different truncations in wave number space.

The first model run includes both synoptic and planetary eddies in the troposphere (the *control* run), while in the subsequent runs synoptic-scale eddies are suppressed (the *truncated* runs). One truncated run includes a mean flow and zonal wave-1 and wave-2, while the other run is truncated to a mean flow and zonal wave-2 only. For

more details about these runs please refer to the description of the model runs in Chapter 2. As a quick reminder, all runs use a γ=4 K/km lapse rate in the equilibrium temperature profile for inducing the stratospheric cold pole, yielding a strong Southern Hemisphere - like polar vortex. The model runs have no seasonal cycle and are run in perpetual winter conditions. For the truncated runs, the only difference from the control run is the described truncation in wave number space. There is no such truncation in the meridional direction. No longitudinally asymmetric forcing is applied in these runs.

4.3 Results

4.3.1 Characteristics of the Differing Model Mean States

As expected, the control run shows characteristics comparable to the Southern Hemisphere: Since no zonally asymmetric surface forcing is applied, we find a lack of quasi-stationary waves. This forcing structure yields a very strong polar vortex exhibiting mean winds around 100 ms^{-1} at its core (Figure 4-7a), with a standard deviation of 5 ms^{-1}. Figure 4-2a shows a representative excerpt from the control run for zonal mean zonal wind $\overline{u}$ at 60°S and 10 hPa (below the jet core). The location for the time series of zonal mean zonal wind was chosen according to the WMO definition of sudden warmings, where a necessary condition for a major sudden warming is defined by a wind reversal at 60° latitude and at a pressure level of 10 hPa.

The corresponding vertical EP flux is shown in Figure 4-2b, depicting latitudinally integrated vertical EP flux at 96 hPa given by $\int_{20^\circ\mathrm{S}}^{70^\circ\mathrm{S}} F_z\, a\cos\varphi\, d\varphi$, where F_z is given by

$$F_z = \rho_0 f a \cos(\varphi) \frac{\overline{v'\theta'}}{\partial\theta/\partial z}, \tag{4.1}$$

where ρ_0 is density as a function of pressure, f is the Coriolis parameter, a is the Earth's radius, φ is latitude, $\overline{v'\theta'}$ is the heat flux and $\partial\theta/\partial z$ is the vertical gradient in potential temperature. This is the quasi-geostrophic spherical definition of the vertical EP flux approximated according to the dominant terms of Equation (1.9). The

EP flux time series is shown at a height of 96 hPa, corresponding to a height where the synoptic tropospheric influence becomes negligible and where the amplitudes of the small wave numbers begin to grow as they propagate into the stratosphere, i.e. the wave flux able to propagate past 96 hPa tends to directly interact with the stratospheric flow.

The time series of the EP flux entering the stratosphere looks noisy (Figure 4-2b). Stratospheric variability is therefore significantly reduced as compared to the real atmosphere with only small bursts of wave activity, which are not able to induce sudden warmings over the entire run.

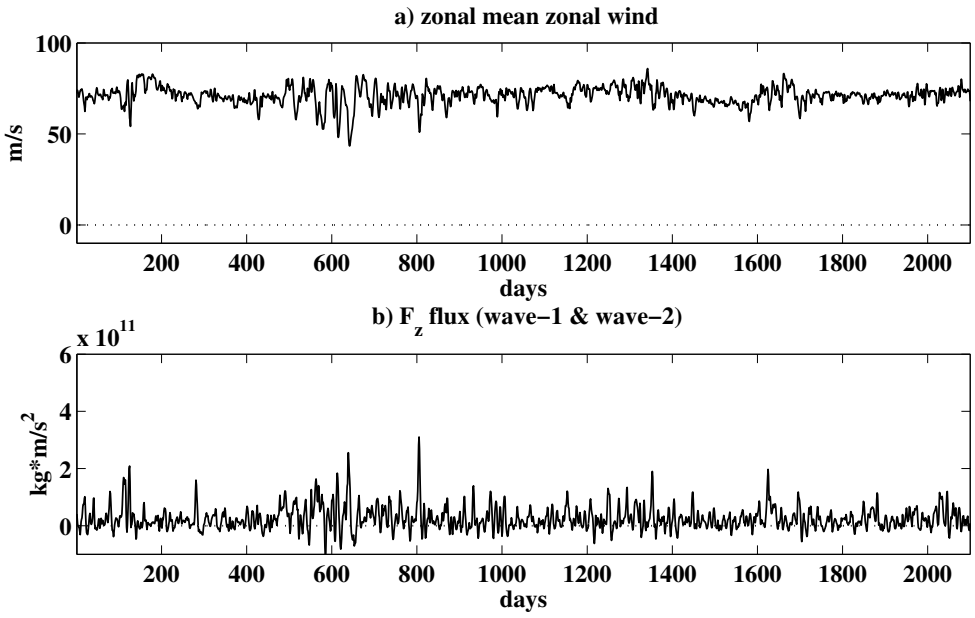

Figure 4-2: Control run: a) Representative part of the time series of zonal mean zonal wind at 60°S and 10 hPa in ms^{-1}. b) Latitudinally integrated vertical component of the EP flux as described in the text for the sum of wave-1 and wave-2 at 96 hPa for the same time period, integrated between 20° and 70°S as given by Equation (4.1).

Both mechanisms described in the introduction to this chapter (the generation of traveling waves by baroclinic instability to long waves as well as by nonlinear interactions among synoptic-scale eddies) are possible candidates for causing the observed traveling waves.

In order to investigate the origin of the traveling waves as well as their impact on

the stratosphere, the control run is compared to two runs where synoptic-scale waves are suppressed. One run suppresses all waves except for a mean flow and wave-2, while the other run allows for a mean flow and both wave-1 and wave-2. Section 2.2.3 provides the mean states of these runs.

The different tropospheric mean states between the described runs dominantly influence the way and magnitude at which tropospheric traveling waves are generated. Comparing Figures 2-3 to Figures 2-6 and 2-5 indicates that the maximum tropospheric wind speeds at the tropospheric jet are shifted equatorward in the truncated runs by several degrees in latitude, and tropospheric westerlies reach further into the tropics in the upper troposphere in the truncated runs. The tropospheric jet is stronger in magnitude in the truncated runs, and all runs show the possibility of a weaker secondary jet centered around 20° latitude. The stratospheric vortex is considerably stronger in the control run, which is due to the increased upward planetary wave flux in the truncated runs as will be discussed in the next section. The stratospheric vortex is more confined to poleward of 30° latitude, while the truncated runs exhibit westerlies further into the tropics.

Suppressing synoptic-scale eddies by a severe spectral truncation not only changes the mean state of the troposphere, but it also yields an increase in planetary wave flux into the stratosphere along with a significant increase in stratospheric variability (Figures 4-3a and 4-4a). In particular, large amplitude warmings occur intermittently associated with large excursions in the EP fluxes (Figures 4-3b and 4-4b). These sudden warmings occur at a similar magnitude and frequency as in the topography run (compare to Chapter 2) for the run truncated to wave-2, but slightly less frequent and weaker in the run truncated to wave-1 and wave-2.

In order to verify that the atmospheric mean states indeed exhibit traveling waves, zonal phase speed spectra for wave-2 for all runs are shown in Figure 4-5. Due to the lack of a surface forcing, there is no forcing of quasi-stationary wave components with phase speeds close to zero. Both traveling wave numbers 1 (not shown) and 2 are present (except of course for the run truncated to wave-2).

All runs exhibit intermittent occurrences of systematic eastward propagation inter-

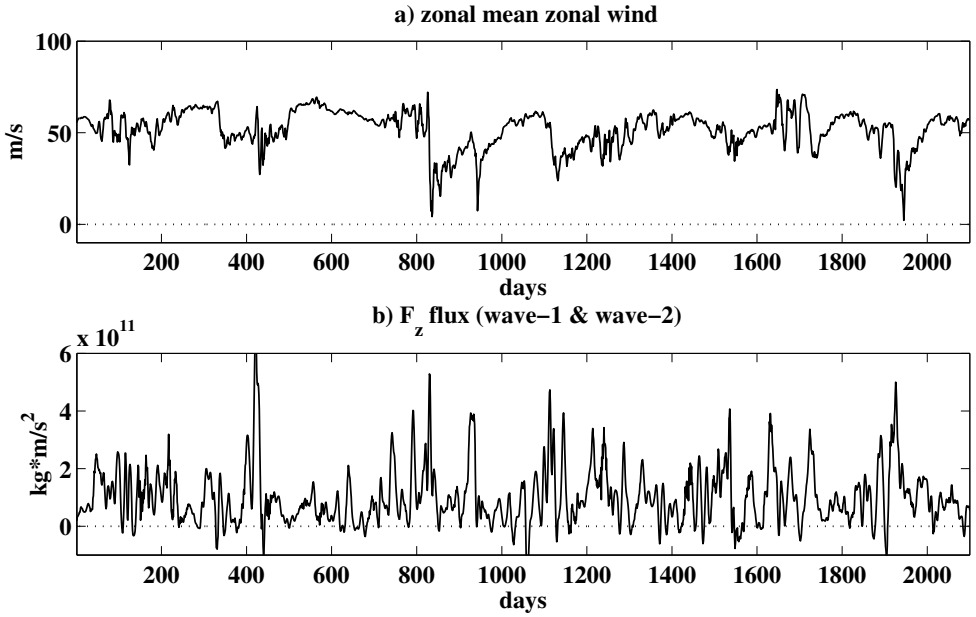

Figure 4-3: Same as Figure 4-2 but for the run truncated to wave-1 and wave-2.

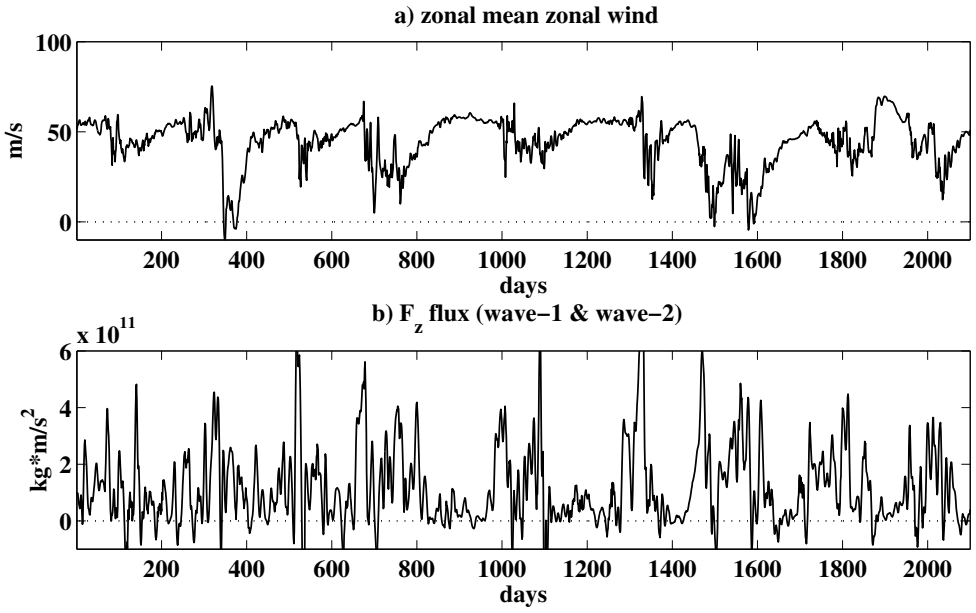

Figure 4-4: Same as Figure 4-2 but for the run truncated to wave-2 only.

spersed with episodes of slower propagation in either direction (Figure 4-5). Wave-2 exhibits maximum eastward phase speeds on the order of 10 - 20 ms^{-1} for both truncated runs, while in the control run, there is broad variability with a less distinct peak at slower eastward phase speeds. The traveling waves are more dominant for the truncated runs, especially for the run truncated to wave-2. A comparison to the real atmosphere can be obtained by e.g. considering the power spectra of the Southern Hemisphere sea level pressure field computed in Mechoso and Hartmann (1982).

As the waves originate in the troposphere, the differing tropospheric mean states and variability between the control and the truncated runs are dominantly responsible for the difference in wave generation. Figure 4-6 shows the growth rate σ as given by

$$\sigma = \left| \frac{f_0}{N} \frac{\partial \overline{u}}{\partial z} \right| \tag{4.2}$$

as a measure of the growth rate of the most unstable mode in the model atmosphere, proportional to the Eady growth rate for the most unstable mode, where $N^2 = \frac{g}{\overline{\theta}} \frac{\partial \overline{\theta}}{\partial z}$ is the Brunt-Väisäla frequency as a measure of atmospheric stability. The derivatives are evaluated in the lower troposphere between the model levels at 514 hPa and 925 hPa.

Figure 4-6 indicates that the most unstable modes exist around the region of the tropospheric jets for all model runs, and the growth rates dominate at the location of their respective jets over the respective other model runs (compare to Figures 2-3, 2-5 and 2-6 for the model mean wind distribution). The mean state of the truncated runs exhibits a secondary peak around 60° latitude, indicating the possibility for wave growth at high latitudes, while for the control run, wave growth maximizes around 45° latitude and falls off strongly poleward of the tropospheric jet. In general, the growth rate increases for the suppression of synoptic-scale eddies.

4.3.2 Analysis of the Generation Mechanism for Traveling Waves

In order to understand the differences in wave generation and propagation between the control and the truncated runs, time-averaged vertical EP fluxes divided into

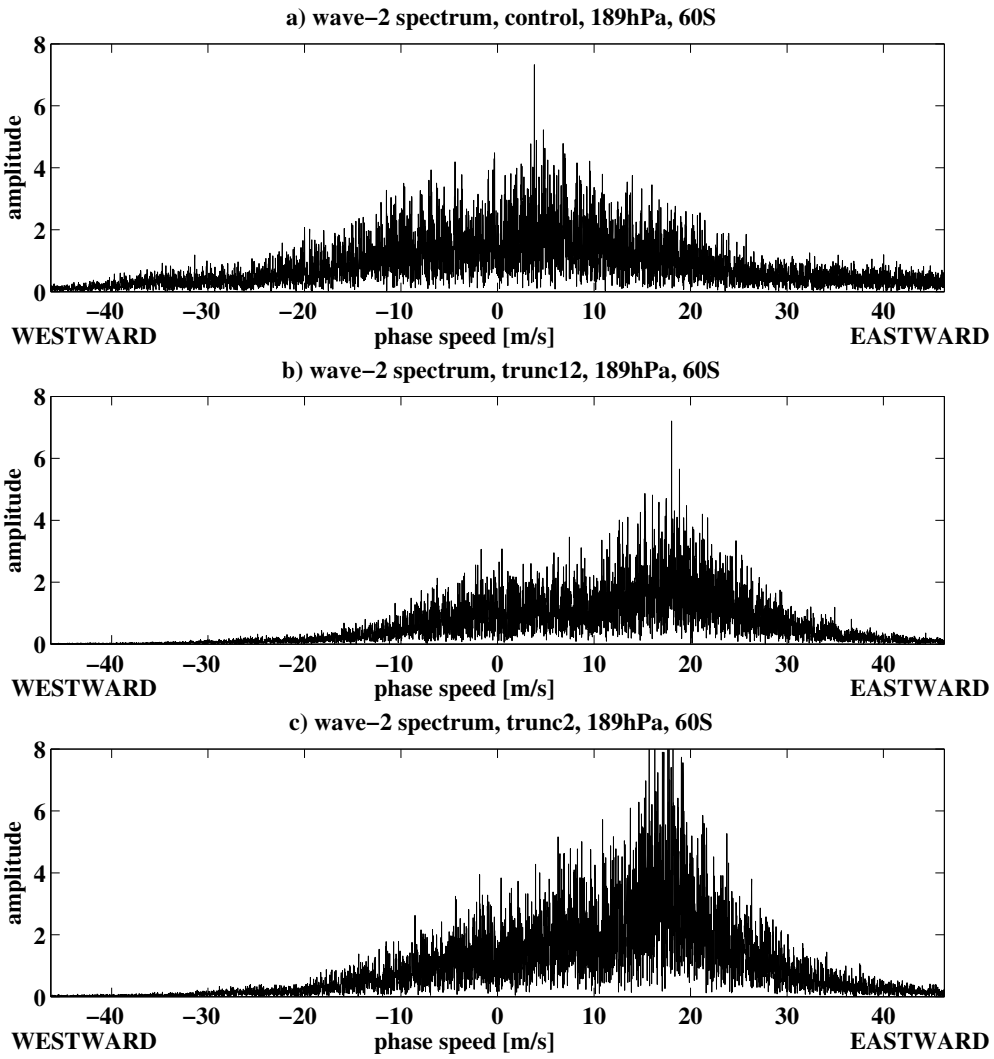

Figure 4-5: Spectrum of the east- and westward phase speeds [ms^{-1}] computed from geopotential height for zonal wave-2 at 189 hPa and 60°S for a) the control run, b) the run truncated to wave-1 and wave-2, c) the run truncated to wave-2. The spectrum was computed using the first 9800 days of each run.

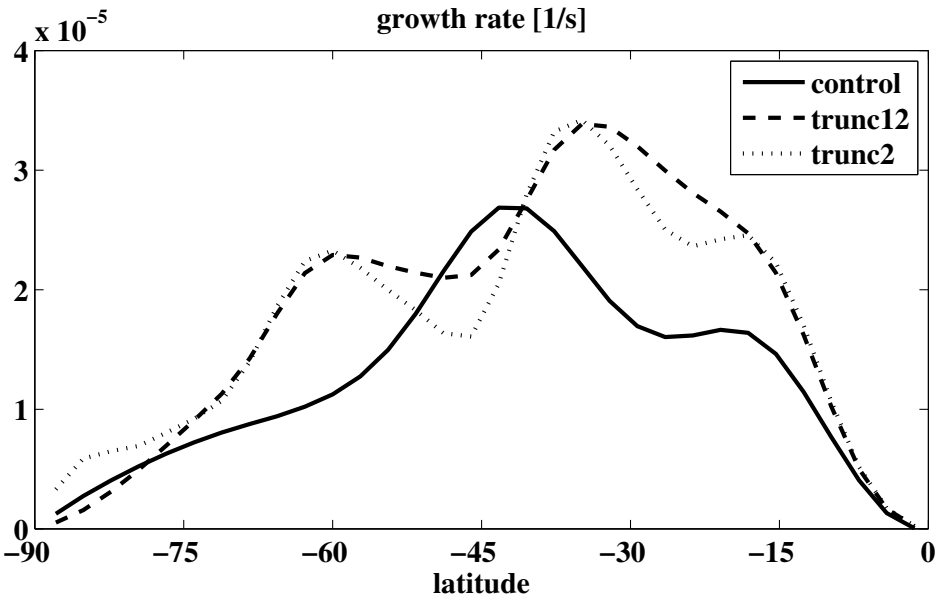

Figure 4-6: Comparison of the growth rates σ [s^{-1}] as a function of latitude for all runs as given by Equation (4.2), which is proportional to the Eady growth rate for the most unstable mode. The solid line is the growth rate for the control run, the dashed line is the growth rate corresponding to the run truncated to zonal wave-1 and wave-2, the dotted line corresponds to the growth rate of the run truncated to wave-2.

planetary-scale (zonal wave-1 and wave-2) and shorter (wave-3 and higher) waves are examined. This separation is based on the finding that wave-3 and higher wave numbers exhibit considerably smaller wave fluxes into the stratosphere as compared to wave-1 and wave-2, and they are not able to significantly alter the stratospheric flow, i.e. every observed sudden stratospheric warming can be linked to the dominant influence of either zonal wave-1, wave-2, or both wave-1 and wave-2.

In the control run the long waves dominate the wave flux in the stratosphere as expected (Figure 4-7b), while the EP flux in the troposphere is dominated by higher wave numbers (Figure 4-7c), with the long waves accounting for only about 10% of the tropospheric wave flux. The origin of the planetary waves in the stratosphere could here be attributed to both or either mechanism mentioned in the introduction.

As the truncated runs inhibit the interaction between synoptic eddies, the mechanism for the generation of traveling planetary waves according to Scinocca and Haynes (1998) is not present, while the mean state of the troposphere has undergone a change

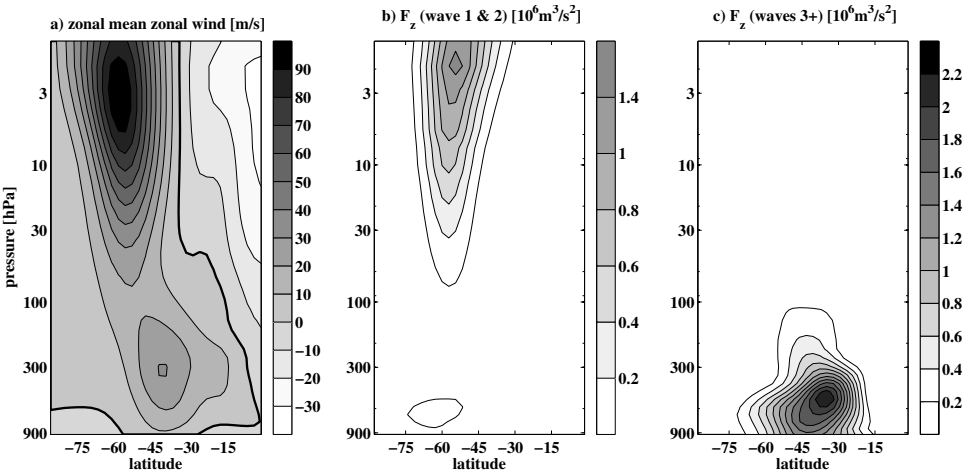

Figure 4-7: Control run: a) Zonal mean zonal wind averaged over the entire run. Contour interval: $10\,\mathrm{ms}^{-1}$. Zero wind line printed in bold. b) Vertical component of the Eliassen-Palm Flux scaled by density ($F_z/\rho_0(z)$), sum of both wave-1 and wave-2. Units in $10^6\mathrm{m}^3\mathrm{s}^{-2}$. Contour interval: $2 \cdot 10^5\,\mathrm{m}^3\mathrm{s}^{-2}$ with contours starting at $2 \cdot 10^5\,\mathrm{m}^3\mathrm{s}^{-2}$. Zero and negative contours omitted for clarity. c) Same as b) but for wave numbers 3 and higher.

due to the absence of synoptic eddies, which has significant implications for the planetary-scale waves which are not truncated. If we would assume that Scinocca and Haynes (1998) was the only mechanism contributing to the generation of planetary-scale waves, we would expect the long-wave flux in the stratosphere to vanish when smaller scale waves are inhibited in the model atmosphere. In the truncated run, however, the long-wave flux strengthens, not only in the stratosphere, but also in the troposphere: Comparing Figure 4-7b to the sum of Figures 4-8b and c as well as Figure 4-9b shows a slight increase in stratospheric EP fluxes in the truncated runs, as well as much stronger tropospheric EP fluxes by wave-1 and wave-2. The strong positive excursions in vertical EP fluxes, which are stronger and more sustained in time than for the control run (comparing Figure 4-2b to Figures 4-3b and 4-4b) are reflected in the higher time-averaged EP flux of the truncated runs. As observed in the real atmosphere, these peaks in wave flux precede strong deviations from the radiatively determined mean zonal winds. The results from the run truncated to zonal wave-2 support the results from the run truncated to wave-1 and wave-2 only, in fact this run exhibits slightly stronger stratospheric variability than the run truncated to

both wave-1 and wave-2.

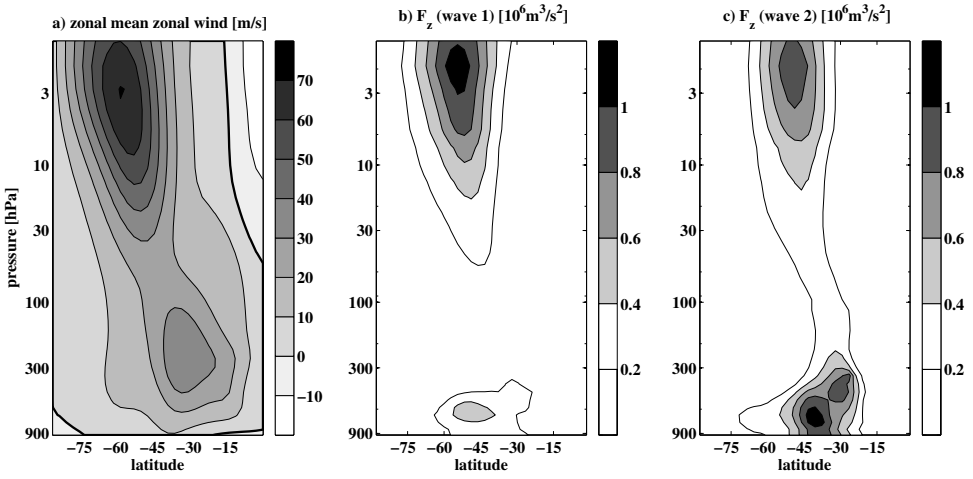

Figure 4-8: Truncated run (wave-1 & wave-2): a) Same as Figure 4-7a. b) Vertical component of the Eliassen-Palm Flux scaled by density ($F_z/\rho_0(z)$), wave-1 only. c) Vertical component of the Eliassen-Palm Flux scaled by density, wave-2 only. Units and contours as in Figure 4-7.

Since in the truncated run, synoptic motions are inhibited, nonlinear interaction between them cannot be responsible for forcing the long waves. Tropospheric baroclinic instability of the long waves themselves, as suggested by Hartmann (1979), must be responsible for the generation of these waves. Instability of the flow to a perturbation by long waves may be enhanced in the truncated run since the suppression of the synoptic-scale waves increases the baroclinicity of the troposphere.

The characteristics of the waves indicate their origin from baroclinic instability: Waves with long zonal wavelengths can become synoptic-scale in terms of their total horizontal wavelength by adopting a large meridional wave number, and by being confined to high latitudes (Hartmann, 1979), and can be classified as Charney modes.

In order to further investigate the meridional length scales of the waves, we perform a comparison between the meridional length scale of zonal wave-2 for the presented runs (Figure 4-10), since wave-2 is present in all runs and as it is the main driver of stratospheric variability in these runs. Assuming the flow is in approximate geostrophic balance,

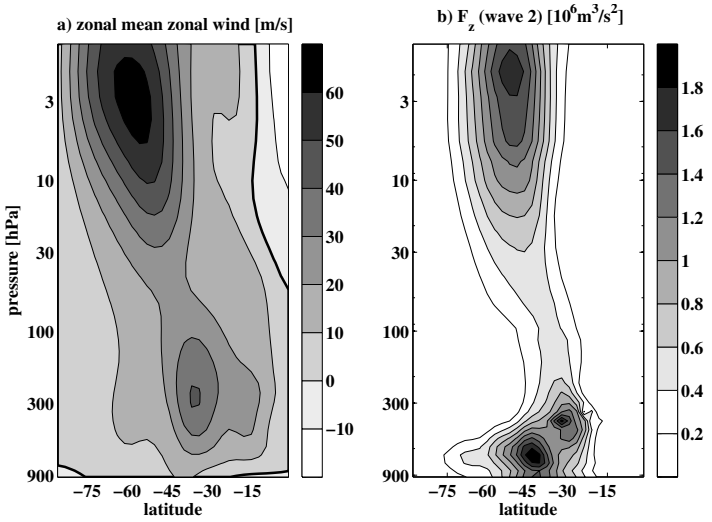

Figure 4-9: Truncated run (wave-2 only): a) Same as Figure 4-7a. b) Vertical component of the Eliassen-Palm Flux scaled by density ($F_z/\rho_0(z)$), wave 2. Units and contours as in Figure 4-7.

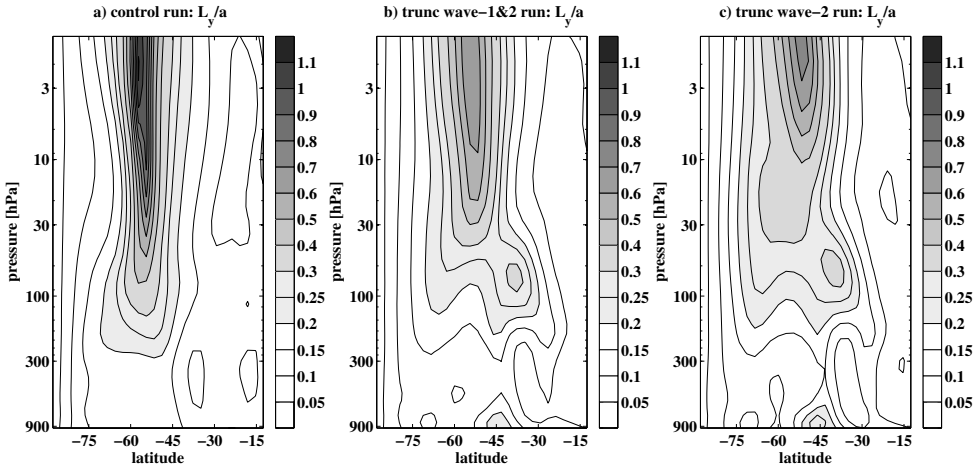

Figure 4-10: L_y/a for zonal wave-2 as a measure of the meridional length scale of the wave in latitude and height as described in Equation (4.3) a) for the control run, b) for the run truncated to wave-1 and wave-2, c) for the run truncated to wave-2. Note the nonuniform contour interval, which is the same for all figures. All figures are averaged in time over the first 10,000 days of the model run, excluding a spin-up period of 100 days.

$$u' \sim \frac{g}{f}\frac{\partial \phi'}{\partial y}\,,$$

where f is the Coriolis parameter, g is the acceleration due to gravity, ϕ' are the geopotential height zonal anomalies and u' are the zonal anomalies of the zonal wind. This then yields

$$< \overline{u'^2} > \sim \frac{g^2}{f^2} < \overline{\frac{\partial \phi'^2}{\partial y^2}} > \sim \frac{g^2}{f^2 L^2} < \overline{\phi'^2} >\,,$$

where an overbar denotes a zonal average and $< . >$ denotes a time average. From this equation we can define a characteristic Southern Hemisphere meridional length scale for zonal wave-2 (indicated by the index in Equation (4.3)), denoted L_y and scaled by Earth's radius a:

$$\frac{L_y}{a} \equiv -\frac{g}{f\,a}\sqrt{\frac{< \overline{\phi_2'^2} >}{< \overline{u_2'^2} >}}\,. \tag{4.3}$$

The minus sign is included in order for $\frac{L_y}{a}$ to be a positive quantity despite looking at the Southern Hemisphere. The quantity $\frac{L_y}{a}$ is of order 1 for planetary meridional scales of the wave, and smaller than 1 for a smaller meridional scale.

Both the control run and the truncated runs indicate a small meridional scale, $L_y \leq 0.2a$, of the wave in the troposphere. The meridional scale increases with height for all runs as required for propagation into the stronger stratospheric winds according to Charney and Drazin (1961).

4.4 Discussion

The above results are here discussed in the light of the mechanism responsible for generating the traveling waves, as well as in terms of stratospheric wave-wave interaction.

4.4.1 Tropospheric Forcing Mechanism for Traveling Waves

In the absence of a stationary planetary-scale tropospheric forcing such as topography or a zonally asymmetric heat pattern, stratospheric variability is dominated by traveling planetary-scale waves which are dominantly generated in the troposphere. The generation of these waves is attributed to different mechanisms including different scales of tropospheric waves and will be considered here.

A comparison of two model runs is shown: a model run truncated to the propagation of planetary-scale waves only is compared to a full model run including both planetary-scale waves as well as synoptic eddies. The truncated run exhibits considerably stronger planetary wave generation and wave flux into the stratosphere. The generation of these waves is attributed to baroclinic instability of the flow to planetary-scale disturbances. The possibility for this instability of the flow to long waves has originally been suggested by Hartmann (1979), and the hypothesis that the observed waves are created by baroclinic instability of long waves is supported by multiple indicators: wave-wave interaction among synoptic-scale waves is inhibited and therefore not responsible for causing the planetary wave flux into the stratosphere, and the small meridional scale of the waves in the troposphere indicates that these modes are produced by tropospheric baroclinic instability.

Adding synoptic eddies opens the possibility for an additional wave generation mechanism, i.e. the one described in Scinocca and Haynes (1998). With another source of planetary-scale wave activity, one may expect that wave generation is enhanced and that the wave flux into the stratosphere increases. However, stratospheric variability as well as planetary-scale wave flux in a spectral core model with no longitudinally varying forcing decreases significantly when synoptic variability is added. So although the truncated model runs exhibit a cleaner structure by only allowing propagation of a single (or two) planetary-scale wave numbers, the stratosphere turns out to be much more variable than for a run also including synoptic-scale wave numbers.

While this finding does not let us answer conclusively what mechanism produces the waves in the control run, it indicates that the mechanism generating the waves in

the truncated runs is damped or suppressed in the control run. Since we suggest that baroclinic instability of long waves causes the planetary waves in the truncated run, it seems that the presence of synoptic eddies damps baroclinic instability of long waves in the control run. This may be explained by the fact that the absence of synoptic eddies in the truncated run strengthens midlatitude baroclinicity, as synoptic eddies are not able to weaken the meridional temperature gradient. The strengthened baroclinicity provides one of the necessary criteria for baroclinic instability of long waves as indicated by Hartmann: through the strengthened baroclinicity, the vertical wind shear is increased through thermal wind balance, contributing to the possibility for baroclinic instability of long waves. There is evidence for a strengthened baroclinicity in the model runs: Comparing Figure 4-7 to Figures 4-8 and 4-9 shows that the lower tropospheric vertical wind shear is strengthened at the location of the tropospheric jet in the truncated runs as compared to the control run. The strengthened baroclinicity in the regions of the tropospheric jets yields an increase in the growth rate of the most unstable mode, as indicated in Figure 4-6. Furthermore, the figure reveals larger growth rates for the suppression of smaller-scale eddies in the truncated run.

This leads to the conclusion that the presence of synoptic eddies does not have a straight forward effect: while on the one hand the synoptic eddies weaken the baroclinicity and thereby lower the possibility for planetary waves to become unstable, thereby reducing planetary-wave generation, they on the other hand represent a possibility for an enhancement of the generation of planetary-scale waves through the Scinocca and Haynes (1998) mechanism.

The question arises however why the wave flux into the stratosphere is so much weaker in the control run. If we conclude that nonlinear interaction between synoptic eddies produces the wave flux in the control run, while baroclinic instability of long waves produces the wave flux in the truncated run, we then also have to conclude that the synoptic eddies are considerably less efficient at producing a steady and strong flux of zonally long waves which are able to propagate into the stratosphere, at least in the given model setting. Even if instability of planetary-scale waves accounts for part of the wave generation in the control run, as may be assumed from the weak

tropospheric planetary-scale wave flux in Figure 4-7, it can still be concluded that the effect of weakening of midlatitude baroclinicity is the dominant effect, while the possibility for the Scinocca and Haynes (1998) mechanism plays a secondary role. Long wave baroclinic instability turns out to be a more efficient driver of stratospheric variability than synoptic eddy interaction producing planetary-scale waves, but only given that baroclinicity is strong enough to support the first mechanism.

While the role of the synoptic eddies it two-fold, it is clear from the results of this study that during times of weak synoptic wave forcing, there is a possibility for the troposphere to become unstable to planetary-scale wave disturbances. In our three-dimensional model atmosphere, these waves are able to propagate and significantly impact the stratospheric flow. This result explains why the flux of wave activity in the Southern Hemisphere winter stratosphere is generally weak: baroclinic instability of synoptic waves efficiently weakens extratropical baroclinicity, while the generation of long waves from synoptic-scale waves organizing into larger-scale wave packets according to the Scinocca and Haynes (1998) mechanism is considerably less efficient as compared to e.g. Northern Hemisphere forcing by topography. This can also be confirmed by comparing the control run to the topography run in Chapter 2. We can therefore conclude that in a typical Southern Hemisphere winter, the Scinocca and Haynes (1998) mechanism generates weak traveling waves which do not significantly impact the stratospheric zonal flow. The truncated run on the other hand may help explain what happens during a period of weak baroclinicity, which may possibly help explain the unexpected amplification of a traveling wave-2 pattern observed in the Southern Hemisphere in 2002.

4.4.2 Sudden Warmings Caused by Atmospheric Internal Variability

Southern Hemisphere winter conditions tend to be unfavorable for strong wave propagation and sudden warmings, as observations of the Southern Hemisphere as well as model studies presented here and elsewhere show. Winter 2002 was very disturbed

long before the warming (Figure 4-1), and Northern Hemisphere stratospheric dynamics show that even for less weakening of the mean flow, precursor structures are common if not necessary for a sudden warming to occur. The precursor structures in winter 2002 may have allowed the waves to propagate through the weakened winds. It is however still a matter of ongoing research to explain the exact causes for that special warming.

Despite the fact that no sudden warmings were observed in the control run, a sudden warming caused by the traveling waves observed in the stratosphere may not be impossible. A similar long (10,000 days) model run by Kushner and Polvani (2005) in a spectral core model with no longitudinally varying forcing and no truncation of synoptic waves exhibited a single sudden stratospheric warming which they found to exhibit similar characteristics to the warming observed in the Southern Hemisphere in 2002. They ascribed the generation of the waves responsible for the major stratospheric warming to the Scinocca and Haynes (1998) mechanism, however the possibility of tropospheric instability to long waves was not considered. In addition, a weaker vortex ($\gamma = 2\,\mathrm{K/km}$ as compared to $\gamma = 4\,\mathrm{K/km}$ in the present run) was used in their run, which yields conditions more favorable for wave propagation. Examining the likelihood of such an extreme event for both vortex strengths $\gamma = 2\,\mathrm{K/km}$, which yielded their warming, to $\gamma = 4\,\mathrm{K/km}$, which was used in the present model run, they find that warm events are exponentially distributed (with their event located at 6 standard deviations), with similar slopes of the distribution for the $\gamma = 2\,\mathrm{K/km}$ and the $\gamma = 4\,\mathrm{K/km}$ event, but with a strong decrease in variance for $\gamma = 4\,\mathrm{K/km}$. Concluding, internal atmospheric variability is able to force extreme events in atmospheric models.

4.4.3 Sudden Warmings Forced by a Single Wave Number

In addition to the above results, a comparison of the two truncated runs gives important insights into the internal dynamics of the stratosphere. The truncated runs differ from each other only in the wave numbers which are allowed to propagate: one run allows propagation of both zonal wave-1 and wave-2, while the other run

allows propagation of wave-2 only. Since these waves amplify and propagate into the stratosphere, it will be elucidating to investigate wave-wave interaction in the stratosphere.

In order to understand if and how a single planetary wave number is able to produce sudden warmings, various studies with simple 1- and 2-dimensional models have looked at the propagation of a single wave number and found that it is able to induce significant variability in a 1- or 2-dimensional model atmosphere. The most prominent example is the study by Holton and Mass (1976), who used a quasi-geostrophic β-plane model including a single wave number and a mean flow as well as a relaxation to a mean temperature profile. They showed that for sufficiently large forcing the stratosphere exhibits strong stratospheric vacillations in time despite the time-independence of the forcing. These vacillations were shown to be caused by the interaction of the wave with the mean flow in the stratosphere.

A similar behavior has been observed in more complex models: A three dimensional dynamical core model run by Scott and Polvani (2004) has shown that propagation of a single planetary wave number into the stratosphere can cause significant stratospheric variability: They force a zonal wave-1 pattern by a time-independent heat source in the lower troposphere, while damping all waves except for wave-1 in the troposphere (below 200 hPa) and relaxing the tropospheric mean flow to a prescribed westerly wind profile in order to ensure the possibility for the wave to propagate. In addition to damping the synoptic eddies, they eliminate the meridional temperature gradient in the troposphere (below 200 hPa), i.e. baroclinic instability as observed in our truncated model run is inhibited. This troposphere is designed merely as a provider of wave-1 flux to the stratosphere, comparable to the Holton and Mass (1976) study. They find that despite the time-independent nature of the forcing, the stratosphere exhibits a highly variable state with an interchange between sudden warmings and a subsequent relaxation to radiative mean conditions. In addition to sudden warmings, Scott and Polvani (2004) also find downward propagation of the stratospheric anomalies into the lower stratosphere. As described in Section 5.4.2, we find the same for the run truncated to wave-2.

The above results indicate that in our model, we may be able to assume that the wave flux provided by baroclinic instability is approximately constant in time, while internal stratospheric dynamics is responsible for the strong time-dependence in the upward wave flux as well as the stratospheric flow variability. This will however be difficult to verify, as the flux entering the stratosphere does not represent the wave flux provided by the troposphere, but rather the amount of wave flux that the stratosphere *allows in*.

Our research confirms and expands the results by Scott and Polvani (2004). They had shown that forcing a single wave number in the troposphere, while damping all other wave numbers in the troposphere, can yield significant stratospheric variability. It is from the Scott and Polvani (2004) run, however, not clear if a single forced wave number could have induced a stratospheric warming without wave-wave interaction between different wave numbers in the stratosphere. Our results show that even when setting all other wave numbers to zero throughout the entire atmosphere, stratospheric variability is significant and sudden warmings are possible.

4.4.4 Wave-Wave Interaction in the Stratosphere

Another important observation can be made about the interaction of planetary-scale waves in the stratosphere. Comparing the EP fluxes between the truncated runs in Figures 4-8 and 4-9 shows that including wave-1 considerably decreases the wave flux of wave-2 into the stratosphere. While wave-2 is by itself responsible for the strong stratospheric variability in the run truncated to wave-2, the wave activity flux of wave-2 is considerably decreased when adding a zonal wave-1 component, along with a decrease in stratospheric variability (e.g. comparing Figure 4-4 to the run truncated to both wave-1 and wave-2 in Figure 4-3).

The upward EP flux for wave-1 more than doubles between the troposphere and the stratosphere (Figure 4-8), while for wave-2 it seems to be the case that not all of the produced waves are able to reach the upper stratosphere for both truncated runs. Assuming that wave-1 is also produced directly by baroclinic instability and is then able to propagate into the stratosphere without much interruption from strong

winds and without being reflected, then it must be assumed that wave-1 reduces the potential for baroclinic instability for wave-2 by weakening baroclinicity for its own generation, while in the run truncated to wave-2 only wave-2 is able to weaken baroclinicity by itself. However, since in both runs not all of the produced wave flux for wave-2 is able to reach the upper stratosphere, it therefore has to be assumed that reflective surfaces may exist for wave-2 which are weakening the net upward EP flux. An interaction between wave-1 and wave-2 is possible as well, possibly producing additional higher wave numbers, however all except for wave-1 and wave-2 will be truncated as specified in the model setup.

Concluding, the addition of planetary-scale wave numbers to the truncated model, as simulated by the run where both wave-1 and wave-2 are allowed to propagate, tends to weaken the wave flux and the ability of the waves to induce stratospheric sudden warmings.

4.5 Conclusion

In summary, it is found that the model troposphere in a run truncated to one (or two) planetary-scale wave numbers is baroclinically unstable to those waves. These waves would have to be damped substantially in the troposphere in order to yield the expected results of a decrease in upward wave flux when truncating the run. Experiments with damped runs (see Section 2.2.4) show that these waves have to be damped at a rate of at least $2\,\mathrm{day}^{-1}$ between the surface and up to 400 hPa or higher in order to yield a stratospheric variability similar to the control run, i.e. Equation (2.9) would have to be modified to damp the waves and the mean flow at different rates as given in Equation (2.19).

These results confirm the possibility for baroclinic instability of planetary-scale Rossby waves in a three dimensional dynamical core model in the absence of synoptic waves, which Hartmann (1979) had proposed and found to be true in a β-plane channel model. While the background basic state which is unstable to planetary-scale wave disturbances is highly artificial, baroclinic instability of long planetary-

scale waves is shown to be much more efficient at producing planetary-scale wave bursts that are able to propagate into the stratosphere as compared to wave-wave interaction between synoptic baroclinic eddies.

It should be noted that the models used in this study as well as the one used in the Hartmann (1979) study are dry models which do not allow for the effects of moisture on tropospheric baroclinic instability. The condensation of water vapor may have a considerable impact on the wave number at which instability of tropospheric waves occurs: Emanuel et al. (1987) find in a two-layer semi-geostrophic model for extratropical conditions that for adding moist processes, the longest waves are destabilized, while the wavelength at which maximum growth occurs decreases slightly. These results are confirmed by Zurita-Gotor (2005). This indicates that in the present model, the addition of moisture may further destabilize the long waves, possibly leading to an enhancement of the mechanism described in Hartmann (1979).

In addition it is here found that a single wave number is able to significantly impact the stratospheric flow, while the addition of other planetary-scale wave numbers tends to weaken both the individual wave flux per wave number as well as the total wave flux into in the stratosphere.

For future model studies of stratospheric variability using a full troposphere as a forcing, model baroclinicity and/or planetary wave propagation have to be damped in order to be able to consider the effect of a single wave number on the stratospheric flow.

It is unclear, however, what role baroclinic instability of long planetary-scale waves plays in the real atmosphere with synoptic eddies present, and it will have to be further investigated if there is a role for baroclinic instability of long waves during times of weak synoptic activity. More complex models may illuminate the role of moisture, while a further exploration of the truncation parameter space will indicate how and when baroclinic instability of long waves becomes important in the extratropical atmosphere.

Chapter 5

Estimating the Angular Momentum Budget for a Sudden Warming Composite from Model Data

5.1 Introduction

During the days around stratospheric sudden warmings, momentum is redistributed in the atmosphere. Planetary-scale Rossby waves are dominantly responsible for the transport of momentum from the troposphere to the stratosphere preceding a stratospheric warming, while the mechanism of momentum exchange following a warming is less straight forward. It has been shown that major stratospheric sudden warmings can cause a significant response in the troposphere for up to two months after an event [e.g. Baldwin and Dunkerton (2001), Polvani and Waugh (2004)], indicating an exchange of angular momentum following a warming. The tropospheric response to stratospheric forcings is often framed in terms of the annular mode index, which experiences a significant change in response to sudden warmings in model composites (Gerber and Polvani, 2009), model ensemble studies (Gerber et al., 2009) as well as

reanalysis composites (Baldwin and Dunkerton, 2001). Since the lower stratospheric response is persistent on monthly timescales and since the tropospheric annular mode characterizes the latitudinal location of the jet stream and is thereby tightly connected to tropospheric weather, the notion of a downward exchange of information strengthens the hopes for an improvement of seasonal weather prediction in the extratropical Northern Hemisphere. However, since the downward influence cannot necessarily be detected for single events, making predictions of the exact tropospheric impact is a more complicated task. In addition, the exact pathway and mechanism of downward influence is still being explored and a topic of current research.

Several theories exist describing possible pathways of downward connection, and several of these will be explored in this chapter. Due to the notion of a possible downward exchange of momentum it can be expected that an analysis of the momentum budget will illuminate the terms responsible for momentum exchange and thereby hint at the mechanism of downward influence. In addition, a momentum budget is expected to elucidate stratosphere - troposphere coupling before a sudden warming.

The following sections will explain the use of the Transformed Eulerian Mean equations for the analysis of momentum exchange, introduce possible mechanisms of momentum transport, derive the momentum budget and analyze the relative importance of the terms. The respective roles of the waves and the induced residual circulation in communicating both the effect of possible precursors as well as the sudden warmings will be investigated. The tropospheric response to both the loss of easterly momentum before the warming through upward wave propagation as well as the observed annular mode response after the warming will be considered. In order to investigate these questions, sudden warming composites from the topography model run (Section 2.2.2) are examined to verify the tropospheric response both before and after the warming. This model run yields a large number of warming events available for analysis and the exact descriptions of friction and stress that go into the model are known, as opposed to for reanalysis data. Composites of sudden warmings use the PC3 criterion (introduced in Chapter 3), capturing strong and deeply penetrating events, in order to focus on events featuring enhanced coupling to the troposphere.

5.2 Mechanisms of Momentum Exchange

Wave propagation and momentum deposition (in the form of a convergence of EP flux) can be observed particularly well in the Northern Hemisphere winter stratosphere, with extreme events of EP flux convergence during stratospheric sudden warmings associated with a weakening of the zonal mean flow. This chapter will investigate the relative contribution of the terms in the TEM momentum Equation (1.3) to a change in zonal mean wind in the atmosphere. In addition the location, timing and magnitude of the redistribution of momentum in the atmosphere will be investigated.

Before a sudden warming happens, it is not clear to what extent the loss of momentum, which is large enough to strongly influence the stratospheric flow, is able to induce a change in the much greater mass of the troposphere. In addition, it is unclear to what extent the waves are amplified within the stratosphere.

After a warming takes place in the stratosphere, it has been shown that the troposphere can be impacted by the change in the stratospheric flow. The question arises if this tropospheric response after the warming is

A) the lingering response to the tropospheric loss of easterly momentum in the lead-up to the warming,

B) a response to a downward influence from the stratosphere, or

C) a tropospheric response to the change in wave propagation properties of the lower stratosphere after a change in stratospheric mean wind distribution.

The dynamics of tropospheric processes suggests dominant tropospheric time scales of 6 - 10 days [e.g. Feldstein (2000)], while mechanism A would require a tropospheric memory of several weeks. Gerber et al. (2009) have further examined the tropospheric response in the light of this mechanism in a three dimensional dynamical core ensemble model run. They perturb the tropospheric flow by small random perturbations to wave numbers 4 to 10 in the winter hemisphere midlatitudes starting 10 days before the stratospheric sudden warming in order to wipe out tropospheric memory. They

then verify the subsequent tropospheric response and find that similarly to Baldwin and Dunkerton (2001), a tropospheric response can be identified in an average over the ensemble members, suggesting that the troposphere responds to a change in stratospheric properties, rather than reacting to a lingering response to tropospheric anomalies.

Options B and C, however, are more difficult to separate. Lower stratospheric anomalies may have a strong influence on the troposphere: The shear (Chen and Robinson, 1992) and magnitude (Charney and Drazin, 1961) of the lower stratospheric winds is of great importance for determining wave propagation into the stratosphere, and there have been suggestions of the lower stratospheric winds influencing upper tropospheric eddy phase speeds (Chen and Held, 2007). This indicates that a tropospheric response could be induced by the state of the lower stratosphere, either passively by influencing upward wave propagation or actively by changing the tropospheric eddy properties.

In general, stratospheric signals are too weak to induce a detectable tropospheric response, and often only an ensemble or a composite of several stratospheric events may show a tropospheric response to a forcing. However, it has been suggested that a significant response can be detected after strong stratospheric forcings such as stratospheric warmings. The strength of the tropospheric response is strongly dependent both on the depth and magnitude of the lower stratospheric response to the forcing in the upper stratosphere as well as on the state of the troposphere at the time of downward propagation of the stratospheric signal (Chan and Plumb, 2009).

5.2.1 Mechanisms of Momentum Transport within the Stratosphere

While lower stratospheric anomalies may have an influence on the tropospheric flow, the question arises how a possible downward momentum exchange from the location of the wave convergence in the upper stratosphere to the lower stratosphere occurs in the first place.

It has both been theoretically predicted (Haynes et al., 1991) as well as observed in both models as well as reanalysis data that the weakening of the mean flow through wave breaking induces a residual circulation in the latitude/height plane which spreads beyond (and in particular below) the forcing location. This mechanism of downward influence has been termed *downward control* after Haynes et al. (1991), or *balanced response*. Downward control can be understood from the TEM Equations (1.3) to (1.7): Following Plumb (1982), Garcia (1987) and Haynes (2005), three of the four independent variables ($\overline{u}$, $\overline{\theta}$, $\overline{v}^*$, $\overline{w}^*$) in the TEM equations can be eliminated to solve for the fourth, e.g. for the vertical part of the residual circulation $\overline{w}^*$, in order to be able to examine the effect of the wave forcing on the residual circulation. This yields an equation for $\overline{w}^*$ governed by an elliptical operator, indicating that the response to a localized wave forcing spreads beyond the forcing location. The spread of the residual circulation is determined by the ratio of the frequency of the forcing to frictional relaxation (which is small or negligible in the stratosphere) and thermal relaxation. Both Haynes et al. (1991) and Holton et al. (1995) have performed idealized numerical experiments estimating the residual circulation response to a localized longitudinally symmetric easterly momentum forcing. For an increase in the forcing frequency relative to the radiative damping rate, the response to the forcing narrows in latitude and strengthens in magnitude, so that in steady state the meridional extent of the circulation is about equal to the meridional extent of the forcing, while extending all the way to the surface where it is balanced by friction. For a comparably short forcing pulse in midlatitudes the response is weaker but can extend into the opposite hemisphere, and an additional residual circulation cell can appear above the forcing region. If the forcing pulse is considerably shorter than the radiative relaxation time scale, the response reduces to the *Eliassen* response [Eliassen (1951), Plumb (1982)].

Another way to determine the downward influence of a sudden warming is the inversion of the induced stratospheric potential vorticity anomaly. Note that a northward eddy flux of potential vorticity is related to the divergence of EP flux observed during sudden warmings, as given under quasi-geostrophic scaling by $\overline{v'q'} = \rho_0^{-1} \nabla \cdot F$.

Hartley et al. (1998) found that the tropopause geopotential height field is significantly impacted by the redistribution of potential vorticity induced by a stratospheric sudden warming as compared to an undisturbed polar vortex.

Critical level mechanics in addition play an important role in the vertical coupling within the stratosphere: Due to the descent of the critical level (where the wind speed matches the phase speed of the propagating waves) the waves tend to break at subsequently lower levels and induce additional residual circulations at lower levels. The critical level may however not descend at the same rate or in the same fashion at all longitudes, complicating the description of this mechanism. Plumb and Semeniuk (2003) found the downward progression of the critical level is solely caused by the interaction of the zonal mean flow with the waves, as opposed to other mechanisms such as wave reflection or the induced meridional circulation.

5.2.2 Momentum Exchange between the Stratosphere and the Troposphere

The above described induced residual circulation is an established signal within the stratosphere which has a strong influence all the way down to the tropopause. However, it is not clear how and if a lower stratospheric or tropopause signal is transferred to the troposphere. Song and Robinson (2004) suggest that the tropospheric response has to be coordinated by synoptic-scale Rossby waves, which organize the response to the stratospheric signal into the intrinsic tropospheric response given by the tropospheric annular modes, a mechanism they termed *downward control with eddy feedback*. Several studies find that the observed change in the surface winds is observed to be and, according to theory, needs to be accompanied by a change in tropospheric eddy momentum fluxes [e.g. Chen and Zurita-Gotor (2008), Chen and Held (2007)] in order to produce the observed annular mode response, and a tropospheric annular mode response can be observed in both observations as well as model simulations of sudden warmings as described in Chapter 1.

On the other hand, Thompson et al. (2006) argue that the balanced response to

anomalous lower stratospheric wave drag is sufficient to induce the observed surface wind anomalies for the longterm effect of ozone depletion in the stratosphere, meaning that the surface winds respond to changes in the induced meridional circulation as opposed to the tropospheric eddy feedback. They simulate the tropospheric response to stratospheric momentum forcing and radiative heating in a zonally symmetric quasi-geostrophic linear model, where the forcings are estimated from observations. Similar forcings can be observed as a response to sudden warmings, although on different time scales. They find that the balanced response to the stratospheric forcing in their model is able to account for the magnitude of the tropospheric response observed in more complex models, while in addition the stratospheric radiative heating increases the persistence of the tropospheric response.

Other mechanisms can be observed to influence the vertical momentum exchange and the connection between the stratosphere and the troposphere. Harnik and Lindzen (2001) find evidence of downward reflection of planetary-scale waves in the winter stratosphere. However, it is not yet resolved how wave reflection events are related to sudden warming events (Shaw and Perlwitz, in press), and the identification of wave reflection events is more involved than the detection of sudden warmings and beyond the scope of this study.

This chapter will examine the exchange of momentum within the atmosphere and between the terms of the TEM equations with the intention to identify the impact of the wave forcing on both the change in mean zonal wind and the induced residual circulation. Ideally, this will hint at the mechanism of momentum exchange between the troposphere and the stratosphere. In particular, the goal is to identify if there is a possibility for a balanced response to anomalous stratospheric wave drag, which would be indicated by the tropospheric residual circulation dominating the response in the troposphere in the absence of meridional momentum fluxes induced by the waves, i.e. F_y. On the other hand, an annular mode response in the troposphere along with significant meridional EP fluxes indicates a smaller role by the meridional residual circulation.

5.3 Derivation of the Angular Momentum Budget

In order to compute the angular momentum budget of the extratropical atmosphere over a zonally averaged domain in time and space, it makes sense to start from the TEM momentum Equation (1.3), which is repeated here for completeness:

$$\frac{\partial \overline{u}}{\partial t} + \overline{v}^*[\frac{1}{a\cos\varphi}\frac{\partial}{\partial\varphi}(\overline{u}\cos\varphi) - f] + \overline{w}^*\frac{\partial \overline{u}}{\partial z} - \overline{X} = \frac{1}{\rho_0\, a\cos\varphi}\,\nabla\cdot F. \tag{5.1}$$

The TEM equations have the advantage of combining all wave-induced forcings into the RHS term $\nabla \cdot F$ in the momentum equation, which makes it easier to anticipate how the zonal mean flow will respond to a wave forcing. For an illustrative analysis of the advantages of the transformed Eulerian mean, see e.g. Dunkerton et al. (1981), Figure 4 and corresponding discussion.

The vertical coordinate used is a log-pressure coordinate $z = -H\ln(\frac{p}{p_s})$ with a reference surface pressure $p_s = 1000\,\mathrm{hPa}$ and a scale height $H = 7\,\mathrm{km}$. Density is then given by $\rho_0(z) = \rho_s\, e^{-z/H} = \frac{p_s}{RT_s}e^{-z/H} = \frac{p(z)}{RT_s} = \frac{p(z)}{Hg}$ where $R = 287\,\frac{\mathrm{J}}{\mathrm{K\,kg}}$ is the gas constant and $g = 9.81\,\frac{\mathrm{m}}{\mathrm{s}^2}$ is the gravitational acceleration. This then yields

$$\frac{\partial p(z)}{\partial z} = gH\,\frac{\partial \rho_0(z)}{\partial z} = -g\rho_s\, e^{-z/H} = -g\rho_0(z). \tag{5.2}$$

5.3.1 The Momentum Budget in Flux Form

For the purpose of numerical integration of the momentum budget, it is useful to put the momentum equation into flux form using the continuity Equation (1.6) in order to simplify the integration of the terms contributing to a change in time of zonal mean wind $\overline{u}$ within the chosen box. All contributions can then be expressed as fluxes across the boundaries of the chosen box. Multiplying Equation (1.3) by $a\cos\varphi$

the advection terms read

$$\begin{aligned}
\overline{v}^* \frac{\partial}{\partial \varphi}(\overline{u}\cos\varphi) + a\cos\varphi\,\overline{w}^* \frac{\partial \overline{u}}{\partial z} &= \overline{v}^* \frac{\partial}{\partial \varphi}(\overline{u}\cos\varphi) + \frac{1}{\rho_0}(\rho_0\, a\cos\varphi\,\overline{w}^* \frac{\partial \overline{u}}{\partial z}) \\
&= \frac{\partial}{\partial \varphi}(\overline{v}^*\overline{u}\cos\varphi) - \overline{v}^*\overline{u}\sin\varphi - \overline{u}\frac{\partial}{\partial \varphi}(\overline{v}^*\cos\varphi) \\
&\quad + \frac{a\cos\varphi}{\rho_0}\frac{\partial}{\partial z}(\rho_0 \overline{w}^*\overline{u}) - \frac{a\cos\varphi}{\rho_0}\,\overline{u}\,\frac{\partial}{\partial z}(\rho_0\overline{w}^*) \\
&= \frac{1}{\cos\varphi}\frac{\partial}{\partial \varphi}(\overline{v}^*\overline{u}\cos^2\varphi) - \overline{u}\frac{\partial}{\partial \varphi}(\overline{v}^*\cos\varphi) \\
&\quad + \frac{a\cos\varphi}{\rho_0}\frac{\partial}{\partial z}(\rho_0 \overline{w}^*\overline{u}) - \frac{a\cos\varphi}{\rho_0}\,\overline{u}\,\frac{\partial}{\partial z}(\rho_0\overline{w}^*)
\end{aligned}$$

using the chain rule. The second and last term on the RHS of this equation add to zero using the continuity Equation (1.6), so that Equation (1.3) with the advection terms in flux form reads

$$a\cos\varphi\frac{\partial \overline{u}}{\partial t} - \overline{v}^* f a\cos\varphi + \frac{1}{\cos\varphi}\frac{\partial}{\partial \varphi}(\overline{v}^*\overline{u}\cos^2\varphi) + \frac{a\cos\varphi}{\rho_0}\frac{\partial}{\partial z}(\rho_0\overline{w}^*\overline{u}) = \frac{1}{\rho_0}\nabla\cdot F + a\cos\varphi\overline{X}.$$

In order to write the Coriolis term $-\overline{v}^* f a\cos\varphi$ in flux form, a streamfunction is defined based on the continuity Equation (1.6):

$$\begin{aligned}
\rho_0 a\cos\varphi\,\overline{v}^* &= -\frac{\partial \chi^*}{\partial z} \\
\rho_0 a\cos\varphi\,\overline{w}^* &= \frac{1}{a}\frac{\partial \chi^*}{\partial \varphi}.
\end{aligned}$$

This yields for the streamfunction as a function of time, pressure and latitude

$$\begin{aligned}
\chi^* &= -\int_{z_{top}}^{z} \rho_0 a\cos\varphi\,\overline{v}^*\,dz \\
&= \int_{p_{top}}^{p} \frac{a\cos\varphi}{g}\,\overline{v}^*\,dp
\end{aligned}$$

using Equation (5.2) and integrating from the top of the model domain (z_{top} or p_{top}) to the chosen vertical level. The streamfunction can equivalently be computed from

$$\chi^* = \int_{\varphi_{pol}}^{\varphi} \rho_0 a^2 \cos\varphi \, \overline{w}^* \, d\varphi$$

integrating to the chosen latitude. The Coriolis term in flux form then reads

$$-\overline{v}^* f a \cos\varphi = \frac{2\Omega \sin\varphi}{\rho_0} \frac{\partial \chi^*}{\partial z}$$

where $f = 2\Omega \sin\varphi$. This yields for the TEM momentum equation in flux form:

$$a\cos\varphi \frac{\partial \overline{u}}{\partial t} + \frac{2\Omega\sin\varphi}{\rho_0}\frac{\partial \chi^*}{\partial z} + \frac{1}{\cos\varphi}\frac{\partial}{\partial\varphi}(\overline{v}^*\overline{u}\cos^2\varphi) + \frac{a\cos\varphi}{\rho_0}\frac{\partial}{\partial z}(\rho_0\overline{w}^*\overline{u}) = \frac{1}{\rho_0}\nabla\cdot F + a\cos\varphi\overline{X}.$$

5.3.2 Derivation of the Lower Boundary Terms

In order to compute the frictional effects, both the form stress of the flow over the topography (if topography is present in the model run) and the drag on the mean flow at the bottom of the model domain have to be taken into account.

Rayleigh drag as described in Chapter 2 is used in order to account for drag at the bottom of the model, following the form of the model Equation (2.15) for the topography run

$$\frac{\partial \overline{u}}{\partial t} = ... + v_{damp}\overline{u}$$

and for the truncated run as given in Equation (2.18)

$$\frac{\partial \overline{u}}{\partial t} = ... + v_{damp}\Big[\overline{u} - \sin^2(2\varphi)\, u_{surf}\Big]$$

where v_{damp} is given by Equation (2.14) (note that v_{damp} is zero or negative) and u_{surf} =5 ms^{-1} for the run truncated to wave-2.

Topographic form stress can be formulated in terms of geopotential as

$$\int_0^{2\pi} \frac{\partial \overline{u}}{\partial t}\, dx = ... - \frac{1}{a\cos\varphi} \int \Big[\frac{\partial \Phi}{\partial x}\Big]_p dx$$

where $[.]_p$ indicates that the derivative is taken along constant pressure surfaces. The RHS will integrate to zero above the mountain.

This then yields for friction $\overline{X}$ in the momentum budget for the topography run

$$\overline{X} = -\frac{1}{2\pi a\cos\varphi} \int \Big[\frac{\partial \Phi}{\partial x}\Big]_p dx + v_{damp}\overline{u}$$

and for the truncated run without topography:

$$\overline{X} = v_{damp}\Big[\overline{u} - \sin^2(2\varphi)\, u_{surf}\Big].$$

These terms will be computed in Section 5.3.4.

5.3.3 Integration of the Budget

This budget can now be integrated over a box A in latitude and pressure weighted by density ρ_0 and $\cos\varphi$. The box is bounded by φ_{equ} on the equatorward side of the box, by φ_{pol} on the poleward side of the box, by z_{bot} at the bottom of the box and by z_{top} at the top of the chosen box, yielding for the integral over the area of the box

$$\int_A [\,.\,]\, \rho_0 \cos\varphi\, dA = \int_{\varphi_{equ}}^{\varphi_{pol}} \int_{z_{bot}}^{z_{top}} [\,.\,]\, \rho_0 \cos\varphi\, dz d\varphi.$$

Integration of the Coriolis term over the box yields

$$\begin{aligned} \int_A \frac{2\Omega\sin\varphi}{\rho_0}\frac{\partial \chi^*}{\partial z}\rho_0\cos\varphi\, dA &= \int_{\varphi_{equ}}^{\varphi_{pol}} \int_{z_{bot}}^{z_{top}} 2\Omega\sin\varphi \frac{\partial \chi^*}{\partial z}\cos\varphi\, dz d\varphi \\ &= 2\Omega \int_{\varphi_{equ}}^{\varphi_{pol}} \chi^*_{top}\, \sin\varphi\cos\varphi\, d\varphi \\ &\quad -2\Omega \int_{\varphi_{equ}}^{\varphi_{pol}} \chi^*_{bot}\, \sin\varphi\cos\varphi\, d\varphi \end{aligned} \tag{5.3}$$

where χ^*_{bot} and χ^*_{top} are the stream function at the bottom and the top of the box, respectively.

Integration of $a\cos\varphi\frac{\partial\overline{u}}{\partial t}$ over the box yields

$$\int_A a\cos\varphi\frac{\partial\overline{u}}{\partial t}\rho_0\cos\varphi\,dA = a\int_{z_{bot}}^{z_{top}}\rho_0\int_{\varphi_{equ}}^{\varphi_{pol}}\cos^2\varphi\frac{\partial\overline{u}}{\partial t}d\varphi\,dz. \tag{5.4}$$

Integration of the flux terms over the box yields

$$\begin{aligned}
&\int_A \frac{1}{\cos\varphi}\frac{\partial}{\partial\varphi}(\overline{v}^*\overline{u}\cos^2\varphi)\,\rho_0\cos\varphi dA\\
=\;&\int_{z_{bot}}^{z_{top}}\rho_0[\cos^2(\varphi_{pol})(\overline{v}^*\overline{u})_{pol}-\cos^2(\varphi_{equ})(\overline{v}^*\overline{u})_{equ}]\,dz\\
=\;&\cos^2(\varphi_{pol})\int_{z_{bot}}^{z_{top}}\rho_0(\overline{v}^*\overline{u})_{pol}dz-\cos^2(\varphi_{equ})\int_{z_{bot}}^{z_{top}}\rho_0(\overline{v}^*\overline{u})_{equ}\,dz
\end{aligned} \tag{5.5}$$

and

$$\begin{aligned}
&\int_A \frac{a\cos\varphi}{\rho_0}\frac{\partial}{\partial z}(\rho_0\overline{w}^*\overline{u})\rho_0\cos\varphi\,dA\\
=\;&\int_{\varphi_{equ}}^{\varphi_{pol}} a\,\cos^2\varphi[\rho_{top}(\overline{w}^*\overline{u})_{top}-\rho_{bot}(\overline{w}^*\overline{u})_{bot}]d\varphi\\
=\;&a\,\rho_{top}\int_{\varphi_{equ}}^{\varphi_{pol}}\cos^2\varphi(\overline{w}^*\overline{u})_{top}\,d\varphi-a\,\rho_{bot}\int_{\varphi_{equ}}^{\varphi_{pol}}\cos^2\varphi(\overline{w}^*\overline{u})_{bot}\,d\varphi
\end{aligned} \tag{5.6}$$

and integration of the EP flux over the box using Equation (1.10) yields

$$\begin{aligned}
&\int_A \frac{1}{\rho_0}\nabla\cdot F\,\rho_0\cos\varphi\,dA\\
=\;&\int_A \frac{1}{\rho_0}\Big[\frac{1}{a\cos\varphi}\frac{\partial}{\partial\varphi}(F^\varphi\cos\varphi)+\frac{\partial}{\partial z}F^z\Big]\,\rho_0\cos\varphi\,dA\\
=\;&\int_A \frac{1}{a}\frac{\partial}{\partial\varphi}(F^\varphi\cos\varphi)\,dA+\int_A\frac{\partial}{\partial z}F^z\cos\varphi\,dA\\
=\;&\frac{\cos(\varphi_{pol})}{a}\int_{z_{bot}}^{z_{top}}F^\varphi_{pol}dz-\frac{\cos(\varphi_{equ})}{a}\int_{z_{bot}}^{z_{top}}F^\varphi_{equ}dz\\
&+\int_{\varphi_{equ}}^{\varphi_{pol}}F^z_{top}\cos\varphi\,d\varphi-\int_{\varphi_{equ}}^{\varphi_{pol}}F^z_{bot}\cos\varphi\,d\varphi.
\end{aligned} \tag{5.7}$$

5.3.4 Integration of the Lower Boundary Terms

For the frictional terms, the following integration can be performed: For the drag on the mean flow, the parameter v_{damp} is given on hybrid levels, corresponding to σ levels at the lowest model levels where the drag is applied, so that the data for $\overline{u}$ and $\overline{p}_s$ are retrieved directly at the hybrid model levels without interpolation to pressure levels, denoted by $\overline{u}^{(\sigma)}$ and $\overline{p}_s^{(\sigma)}$, while $v_{damp}^{(\sigma)}$ on the σ levels close to the surface is given by Equation (2.14). For the topography run, integration of the Rayleigh damping term yields

$$\begin{aligned}
\int_A v_{damp}^{(\sigma)}\,\overline{u}^{(\sigma)}\,a\rho_0\cos^2\varphi\,dA &= a\int_{\varphi_{equ}}^{\varphi_{pol}}\int_{z_{bot}}^{z_{top}} v_{damp}^{(\sigma)}\,\overline{u}^{(\sigma)}\,\rho_0\cos^2\varphi\,dz\,d\varphi \\
&= -\frac{a}{g}\int_{\varphi_{equ}}^{\varphi_{pol}}\int_{p_{bot}}^{p_{top}} v_{damp}^{(\sigma)}\,\overline{u}^{(\sigma)}\,\cos^2\varphi\,dp\,d\varphi \\
&= -\frac{a}{g}\int_{\varphi_{equ}}^{\varphi_{pol}}\cos^2\varphi\int_{\sigma_{bot}}^{\sigma_{top}} v_{damp}^{(\sigma)}\,\overline{u^{(\sigma)}\,p_s^{(\sigma)}}\,d\sigma d\varphi \quad (5.8)
\end{aligned}$$

using Equation (5.2) and $d\sigma = \frac{dp}{p_s}$ with $\overline{p}_s^{(\sigma)}$ the zonal mean surface pressure on the surface σ level, where $\sigma = \frac{p}{p_s}$ is valid for the bottom model levels where the damping is applied, below the transition to hybrid levels. This can be derived from Equation (2.1) with $A_k = 0$ for the bottom levels and B_k corresponding to the σ levels. This integral is zero if $\sigma_{bot} < 0.7$ since $v_{damp}^{(\sigma)}(\sigma < 0.7) = 0$.

For the truncated run the corresponding integration yields

$$\begin{aligned}
\int_A v_{damp}\Big[\overline{u}^{(\sigma)} &- \sin^2(2\varphi)\,u_{surf}\Big]\,a\rho_0\cos^2\varphi\,dA \\
&= -\frac{a}{g}\int_{\varphi_{equ}}^{\varphi_{pol}}\cos^2\varphi\int_{\sigma_{bot}}^{\sigma_{top}} v_{damp}^{(\sigma)}\,\overline{u^{(\sigma)}p_s^{(\sigma)}}\,d\sigma d\varphi \\
&\quad +\frac{a}{g}\int_{\varphi_{equ}}^{\varphi_{pol}}\cos^2\varphi\int_{\sigma_{bot}}^{\sigma_{top}} v_{damp}^{(\sigma)}\sin^2(2\varphi)\,u_{surf}\,\overline{p_s^{(\sigma)}}\,d\sigma d\varphi.
\end{aligned}$$

The form stress due to topography in the topography run can be integrated according to

$$
\begin{aligned}
-\frac{1}{2\pi}\int_A\int_x\left[\frac{\partial\Phi}{\partial x}\right]_p dx\,\rho_0\cos\varphi\,dA &= -\frac{1}{2\pi}\int_\varphi\int_z\int_x\left[\frac{\partial\Phi}{\partial x}\right]_p dx\,\rho_0\cos\varphi\,dz\,d\varphi \\
&= \frac{1}{2\pi}\int_\varphi\frac{\cos\varphi}{g}\int_p\int_x\left[\frac{\partial\Phi}{\partial x}\right]_p dx\,dp\,d\varphi \\
&= \frac{1}{2\pi}\int_\varphi\cos\varphi\int_p\int_x\left[\frac{\partial Z}{\partial x}\right]_p dx\,dp\,d\varphi,
\end{aligned}
$$

where geopotential height is given by $Z=\frac{\Phi}{g}$. The inner integral can now be converted according to

$$
\begin{aligned}
\int_p\int_x\left[\frac{\partial Z}{\partial x}\right]_p dx\,dp &= \int_{p_s}^{p_{top}}\int_x\left[\frac{\partial Z}{\partial x}\right]_p dx\,dp \\
&= \int_{p_s}^{p_*}\int_{x_W}^{x_E}\left[\frac{\partial Z}{\partial x}\right]_p dx\,dp \\
&= \int_{p_s}^{p_*}2\left[Z_E-Z_W\right]dp \\
&= \int_S Z_s^{(\sigma)}\frac{\partial p_s^{(\sigma)}}{\partial x}\,dx\,,
\end{aligned}
$$

where p_* is the first level above the topography, above which the integral yields zero, and p_s is surface pressure as a function of space and time. Here, $Z_s^{(\sigma)}$ denotes the geopotential height on the surface σ level. Z_E and Z_W denote the points where the pressure level intersects with the mountain on the western and eastern side of each of the two valleys of the zonal wave-2 topography (compare to Equation (2.17)). The last equal sign is valid since the integration along both the zonal direction x and pressure p is equivalent to an integral along the topography surface, denoted as S. Since the integral is along the topography, p_s is evaluated directly at the lowest model σ level, while Z_s is given by Equation (2.17) or by geopotential height at the surface σ level, yielding the same result. This then yields for the total form stress due to topography

$$
-\frac{1}{2\pi}\int_A\int_x\left[\frac{\partial\Phi}{\partial x}\right]_p dx\,\rho_0\cos\varphi\,dA = \frac{1}{2\pi}\int_\varphi\cos\varphi\int_S Z_s^{(\sigma)}\frac{\partial p_s^{(\sigma)}}{\partial x}\,dx\,d\varphi. \tag{5.9}
$$

All of the above terms are now integrated numerically to yield the angular momentum budget for any chosen part of the atmosphere.

5.4 Results

5.4.1 Analysis of the Annular Mode Response to Stratospheric Warmings

The notion of stratospheric wave convergence being followed by tropospheric anomalies can be illustrated by same-signed annular mode anomalies in the stratosphere and the troposphere, shown in Figure 5-1 for the topography run (compare Chapter 2). The annular mode index is here computed separately at every vertical level from the principal component time series of the dominant EOF mode of zonal mean zonal wind for the winter hemisphere, following Baldwin and Dunkerton (2001) and Gerber and Polvani (2009). The composite is computed using the PC3 criterion (introduced in Chapter 3). Figure 5-2 shows the composite of zonal mean zonal wind for the same composite criterion (Figures 3-4 and 3-5 show the same composites but for the WMO and PC2 criteria).

The structure of the dominant mode at a representative tropospheric and a stratospheric level indicate the dominant variability in the extratropical atmosphere (Figure 5-3). The stratospheric mode exhibits a maximum at the location of the polar vortex, indicating a cycle of weakening and strengthening of the polar vortex. The tropospheric mode exhibits a node at the location of the tropospheric jet core, indicating that a strengthening at the poleward flank of the tropospheric jet coincides with a weakening at the equatorward side of the jet, indicating a wobbling of the tropospheric jet about its latitudinal mean position. Comparing the signs of the stratospheric and the tropospheric modes indicates that a strengthening of the polar vortex goes along with a poleward displacement of the tropospheric jet (defined by convention as a *positive annular mode index*). The opposite applies as well: A weakening of the polar vortex, as e.g. observed during stratospheric warmings, is expected to correspond

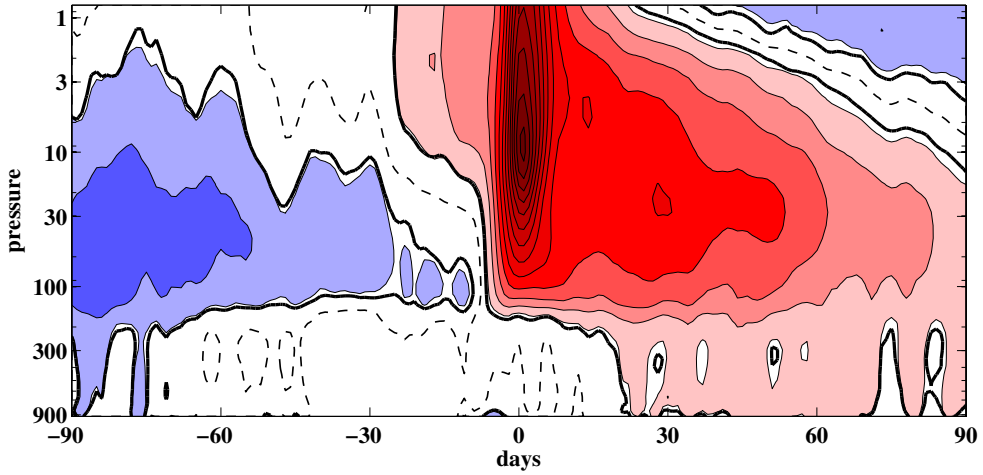

Figure 5-1: Stratospheric Sudden Warming composite of the principal component time series of level-by-level EOF 1 computed from zonal mean zonal wind for the topography run. The sudden warmings are identified using the PC3 criterion, yielding 55 warming events. The bold line indicates the 95% confidence interval, while the dashed line indicates the zero-line between the positive and negative patterns. Red shading indicates a negative annular mode index corresponding to a weak polar vortex and an equatorward shift in the tropospheric jet, while a blue shading corresponds to the opposite pattern.

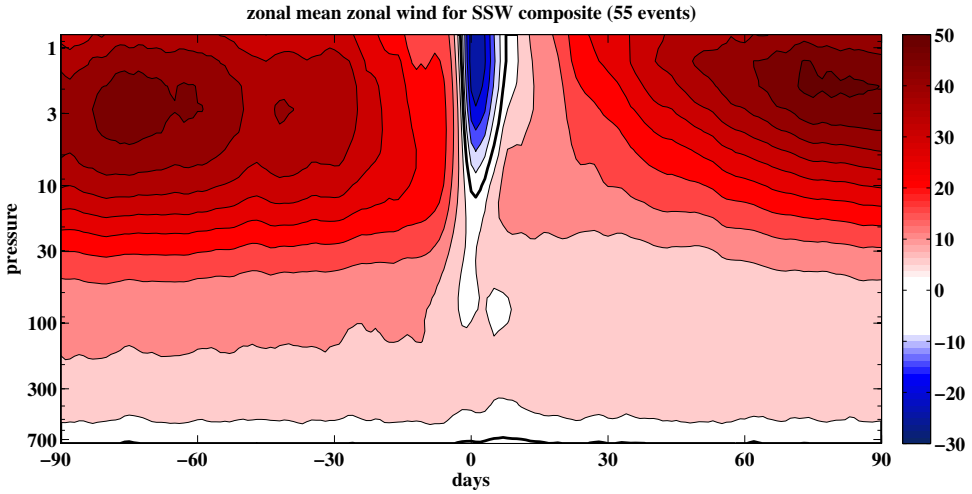

Figure 5-2: Topography run composite time series of daily zonal mean zonal wind [ms^{-1}] at 60°S of all sudden warming events based on the PC3 criterion yielding 55 events. Day 0 corresponds to the day when the criterion is fulfilled. The black solid line is the zero-wind line. Contour interval: $5\,\mathrm{ms}^{-1}$. This composite can be compared to Figures 3-4 and 3-5 which show the corresponding composites for the WMO and PC2 criteria, respectively.

to an equatorward shift of the tropospheric jet (defined by convention as a *negative annular mode index*).

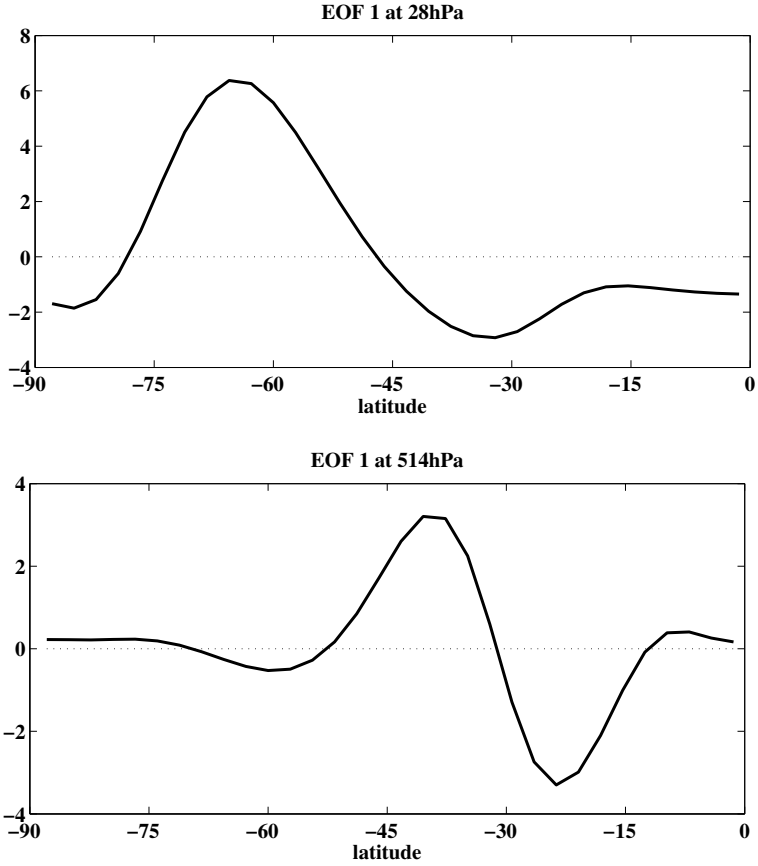

Figure 5-3: Dominant EOF mode of level-by-level EOFs as a function of latitude with units of standard deviation for a representative stratospheric (top) and tropospheric level (bottom), respectively. The top panel shows the dominant EOF structure at 28 hPa, while the bottom panel shows the EOF at 514 hPa as a function of latitude.

The annular mode composite (Figure 5-2) indicates that a significant response can be observed throughout the stratosphere within days after the warming. This indicates that the mechanism involved in bringing the signal to the lower stratosphere acts on time scales of a few days. Considering the mechanisms introduced in Section 5.2.1, any of the mentioned mechanisms or a combination of them may contribute to the downward propagation of the upper stratospheric signal, i.e. the induced residual circulation, wave-mean flow interaction yielding a descent of the critical line, or planetary wave reflection. Most likely their effects are connected: the critical line descends into the mid-stratosphere on time scales of days, thereby lowering the level of

wave breaking and mixing, and possibly even changing the mean flow characteristics to induce reflective layers. While it is not clear how far down the influence of the critical line yields, in addition, a residual circulation is set up which penetrates to the lower stratosphere. The descent of the critical line may however be able to set up additional residual circulations further down, responding to the change in the wave breaking region. The lower stratospheric response to the initial forcing in the upper stratosphere may therefore be a combination of the above mentioned effects.

Once a lower stratospheric signal is established, the downward propagation of the signal comes to a halt, and the signal does not immediately start to descend into the troposphere. While same-signed anomalies of the annular mode index can be observed shortly before and after the warming, they do not become significant until about a month after the stratospheric warming. Changing the significance level from 95% to 99% does not change this feature of the composite. This time scale also corresponds to the decorrelation time scale of the dominant tropospheric EOF mode, which is on the order of 35 days in the upper troposphere, indicating that this time scale may correspond to the adjustment of the eddy momentum fluxes to the lower stratospheric forcing. The observed modes in Figure 5-3 support this notion, as the tropospheric mode corresponds to the latitudinal wobbling of the tropospheric jet whose location is determined by the tropospheric eddy momentum fluxes.

According to Thompson et al. (2006), a residual circulation is expected to be induced by the signal in the lower stratosphere once it is established. An additional residual circulation downward from the lower stratosphere and penetrating the troposphere would then be expected.

5.4.2 Momentum Budget Analysis of Stratospheric Sudden Warmings

In order to examine the indicated exchange of relative angular momentum between the stratosphere and the troposphere, the momentum budget derived earlier in this chapter is now employed, with the goal of a clarification of the responsible mechanism.

The budget of relative angular momentum within boxes each representing part of the winter hemisphere is considered. The momentum fluxes across the boundaries of each box are computed and compared to the change of angular momentum within the respective box. The change of momentum within a given box is represented by the temporal change in mean zonal wind integrated over the box given by Equation (5.4). The vertical EP flux (F_z) through the top and bottom boundaries of the box as well as the meridional EP flux (F_y) through the lateral boundaries of the box are given by Equation (5.7). The EP fluxes act to weaken the mean flow within the box if they enter the box, and strengthen the flow if they leave through the boundaries of the box. The contribution from the Coriolis term at the bottom and top boundaries of the box are given by the two terms in Equation (5.3), while the remainder of the vertical residual circulation through the top and bottom boundaries is given by Equation (5.6). The flux of the residual circulation through the lateral boundaries at the poleward and equatorward boundaries of the box is given by Equation (5.5). Residual circulation terms that enter the box tend to strengthen the mean flow. It can be shown that the Coriolis term dominates the residual circulation, while the remaining residual fluxes have negligible effects on the momentum budget. For a box with a lower boundary close to the surface, the surface drag terms have to be included in terms of topographic form stress given by Equation (5.9) and Rayleigh damping on the mean flow applied in the surface layers as given by Equation (5.8).

In order to study relative momentum exchange, the extratropical atmosphere is divided into three areas in the latitude / height plane: A tropospheric box, a lower stratospheric box, and an upper stratospheric box. The lower and upper boundaries of the boxes are chosen at 1, 10, 96 and 514 hPa, with lateral boundaries at 88 and 32°S as illustrated by the green boxes in Figure 5-4. The equatorward boundary was chosen to lie on the node of the tropospheric annular mode structure (compare Figure 5-3) in order to be able to identify the characteristic latitudinal variation of the tropospheric jet indicated by the annular modes. The tropospheric box will therefore be able to identify an induced annular mode signal as a response to the loss of easterly momentum to the stratosphere before the warming, as well as in

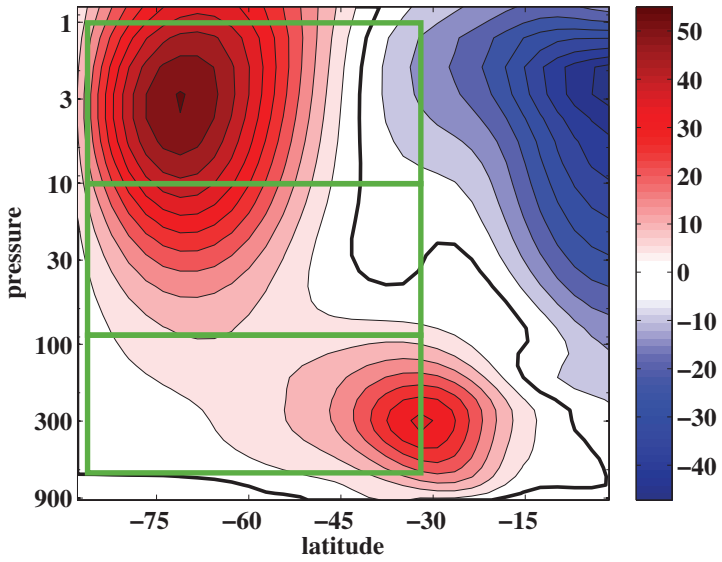

Figure 5-4: Schematic illustration of the boxes for the computation for the angular momentum budget with boundaries at 1, 10, 96 and 514 hPa and lateral boundaries at 88 and 32°S. The boxes are plotted onto the zonal mean zonal wind for the topography run.

response to the stratospheric forcing given by the stratospheric warming. The lower stratospheric box will identify the transport of momentum between the troposphere and the upper stratosphere, while the upper stratospheric box encompasses the polar vortex and captures the effect of the sudden warming in the stratosphere.

Stratospheric Response to Sudden Warmings in the Momentum Budget

This section will focus on the response of the stratospheric momentum budget to the wave forcing from the troposphere. The lower stratospheric box is bounded at the bottom at 96 hPa and by 10 hPa at the top, and by lateral boundaries at 32°S and 88°S, while the upper stratospheric box is bounded at the bottom by 10 hPa and by 1 hPa at the top, with the same lateral boundaries as the lower stratospheric box. Figure 5-5 shows the budget over the two stratospheric boxes in terms of the contributions of the dominant terms: the Coriolis term (blue), the vertical EP flux component (red), and the meridional EP flux component (gray). The flux at the lower box boundary corresponds to the bold line, while the flux at the upper boundary of

the box is represented by the thin line. All terms except for the term including $\frac{\partial u}{\partial t}$ are anomalies with respect to the longterm mean of the respective terms averaged over the entire model run. The variability of these terms which is observed before and after the time of the sudden warming (i.e. before day -10 and after day 30) is characteristic of the variability observed throughout the model run.

For the upper stratospheric box, the fluxes through the bottom of the box for both F_z as well as the Coriolis term by far exceed the fluxes through the top, so that the net contributions are dominated by the fluxes through the bottom boundary. For the lower stratospheric box, the fluxes at the lower and upper box boundaries are comparable in magnitude. As a reminder, negative contributions to the budget correspond to EP fluxes entering the box or the residual circulation term leaving the box. For the meridional EP flux, as expected the influx at the polar latitudes is negligible, so that the F_y budget is dominated by the flux leaving the box at 32°S. The incoming EP flux at the bottom boundary shows comparably strong and slightly increasing values over the days -30 to -10 of the warming composite, followed by a strong convergence into the vortex region leading to the stratospheric warming. The strong growth of the upward EP fluxes starts several days earlier in the lower box as expected from upward wave propagation. At the same time, the Coriolis term balances the incoming F_z flux almost instantly, although it does not exactly match the magnitude of the increase in wave flux. This imbalance is partly balanced by the outgoing F_y flux at the equatorial boundary, especially in the upper box, while the remaining residual leads to a weakening of the mean flow ($\frac{\partial \overline{u}}{\partial t} < 0$) as expected for a sudden warming, indicating a deposition of momentum in the upper stratospheric box.

While the impact of the wave flux in the polar stratosphere is well balanced on daily timescales (but not exactly in magnitude) by the Coriolis term, one would not expect much upward wave flux further equatorward. However, as predicted by the numerical simulations of Holton et al. (1995), a remote effect can be expected in the residual circulation further equatorward as described in Section 5.2.1. For a comparably short forcing pulse such as the short term increase in EP flux observed in Figure

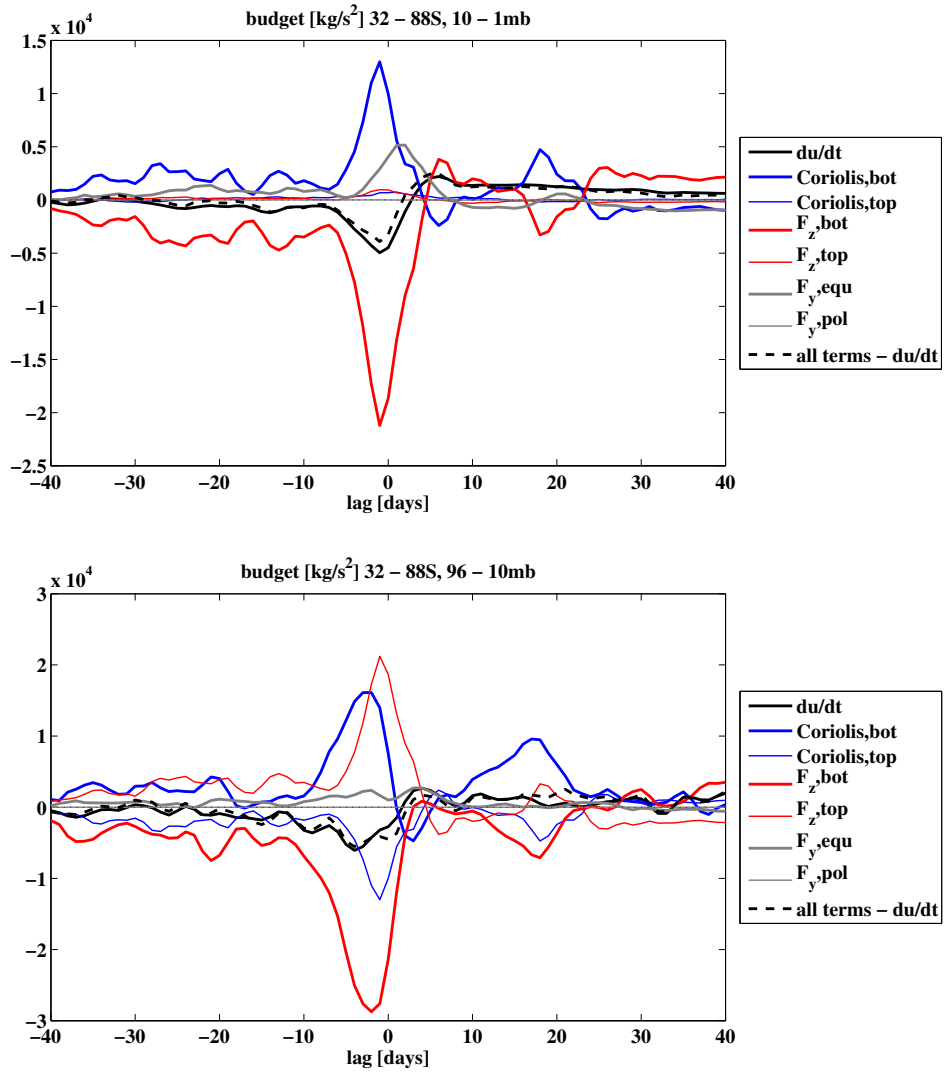

Figure 5-5: Budget of the upper stratosphere (top) and the lower stratosphere (bottom) as a function of lag [days before and after the warming, composited according to the PC3 criterion]. Units: $\mathrm{kg\,s^{-2}}$. The heavy lines correspond to the Coriolis term (blue) and the vertical EP flux (F_z) (red) at the bottom of the respective box, and the meridional EP flux (F_y) (gray) on the equatorward boundary of the box. The thin lines correspond to the same quantities but at the respective opposite side of the box where the terms are smaller. All terms are given as anomalies with respect to the longterm mean of these quantities over the entire model run. The black line denotes the (absolute value of the) net momentum change within the box. The remaining contributions of the residual circulation are not plotted as they are small. The dashed black line is the sum of all terms except for the net change within the box.

5-5, an upward residual circulation is expected in the tropical stratosphere. Figure

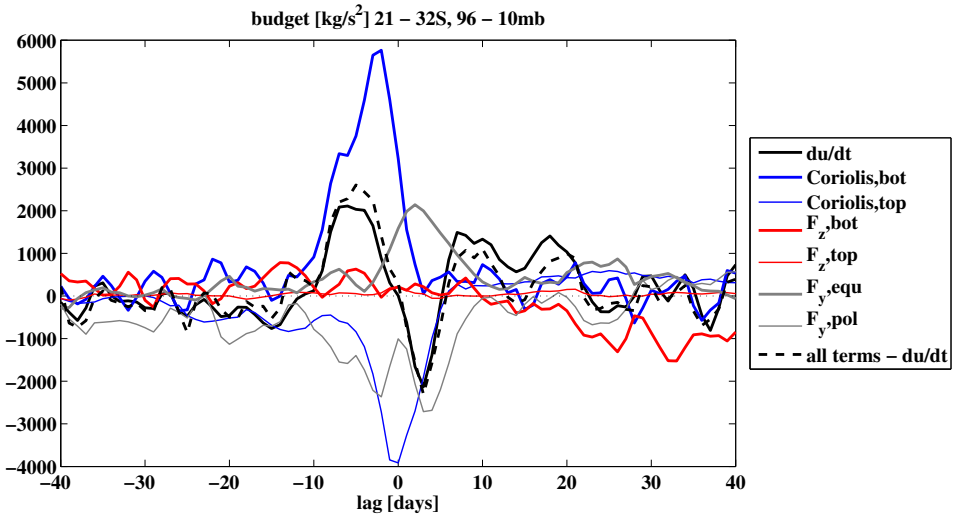

Figure 5-6: Same as Figure 5-5 but for the equatorial lower stratosphere. The box is bounded at 96 hPa at the bottom and 10 hPa at the top, and by lateral boundaries at 21° and 32°S.

5-6 shows the budget for the lower stratosphere, equatorward of the boxes shown in Figure 5-5, between 96 hPa and 10 hPa, and bounded in latitude at 21° and 32°S. While F_z exhibits little variability throughout the composite as compared to the other terms, the Coriolis term mimics the burst observed in the extratropical stratosphere over the week before the warming. This can be described as the remote effect of the wave forcing in the polar stratosphere as expected from Holton et al. (1995). Since F_z and the Coriolis term are not balanced outside of the extratropical stratosphere, the wind exhibits the opposite response to the polar stratosphere by increasing along with the Coriolis term at the bottom boundary before the warming, followed by a sharp decay at the time of the warming. After the stratospheric warming, the wind in the extratropical stratosphere relaxes and starts recovering, an effect which is balanced by the opposite effect in the tropics.

A Note on the Compositing of the Sudden Warmings in the Momentum Budget

A secondary increase in F_z peaking around day 18 can clearly be observed in the budget for the extratropical stratosphere. A first idea would be for the peak to be due to the compositing of the sudden warmings according to the PC3 criterion used for this composite, which picks up events as soon as the principal component time series falls below 3 standard deviations, as opposed to when the principal component reaches a local minimum. Major warmings are often preceded by minor warmings and similar precursor structures as described in Chapter 3. The PC3 criterion picks up several of the warmings before the minimum is reached, i.e. by picking up a strong precursor structure. This can be confirmed by relaxing the criterion to PC2, which yields a much stronger second peak. In addition, a typical time scale for precursor structures to occur is around 2 to 3 weeks before the major warming, indicating that the secondary peak at day 18 is an indication of an additional weakening of the mean flow. Further strengthening of the criterion beyond PC3 was employed in an attempt to avoid the second peak, however a strong decrease in the number of warmings included in the composite is observed, thereby significantly weakening the statistics of the composite.

In order to further investigate if the secondary peak is due to the compositing of the sudden warmings, an additional compositing method was employed. This criterion picks up events when the principal component time series lies below 3 standard deviations in addition to exhibiting a local minimum over a range of 81 days (+/- 40 days from the local minimum) of the time series. This method yields 53 warmings (as compared to 55 warmings in the PC3 criterion). These warmings correspond closely to the warmings identified in the PC3 criterion, while a few of them are identified a couple of days later than when using the PC3 criterion. However, it turns out that this method only marginally changes the budget, and the secondary peak persists. It is therefore likely that the reason for the secondary peak lies in the way the principal component picks up events in the first place. An event characterized by

e.g. a stronger deceleration in wind as compared to a second event may be picked up as weaker in the principal component time series, yielding that several events which would be picked up as major warming using the WMO criterion are characterized as minor in the principal component time series. However, the WMO criterion is less practical to use as discussed in Section 3.4.1 and therefore PC3 will be used throughout this section. This can be justified by the fact that the major warming is throughout most of the time series picked up at a minimum of the principal component time series. In addition it is reasonable to expect that strong precursor structures have a considerable effect on the lower stratosphere as well, indicating that this way of compositing will not weaken the analysis.

Tropospheric Response to Sudden Warmings in the Momentum Budget

The tropospheric response to a stratospheric forcing is more delicate to investigate. During the period leading up to a stratospheric warming, the troposphere loses easterly momentum to the stratosphere. One could expect that this loss may influence the tropospheric flow, manifested in a change in the strength of the tropospheric zonal mean zonal wind. However, due to the greater mass of the troposphere, responses to the same forcing are smaller in the troposphere and therefore difficult to identify against the background of internal tropospheric variability. The effect of mass on the expected change in wind speed can be estimated from the relation of momentum and wind speed by the effect of the mass which is moved by the applied momentum, e.g. the same applied momentum which induces a change of $50\,\mathrm{ms}^{-1}$ in the stratosphere at and above 10 hPa, if applied in the mid-troposphere at 500 hPa and below, will induce a change of only $1\,\mathrm{ms}^{-1}$. The question, however, is how the effect of the applied momentum is transferred between these layers. This section will address the transfer of momentum during the periods both before and after a sudden warming.

Figure 5-7 shows the momentum budget for the upper tropospheric box bounded at 514 hPa at the bottom and 96 hPa at the top, and again bounded by lateral boundaries at 88° and 32°S (like the stratospheric budgets). A clear and almost instantaneous balance of F_z and the Coriolis term can be observed, with F_y contributing

considerably less to the budget. The change in mean wind is a small residual of the large balanced fluxes of waves and the residual circulation. Again, the budget is dominated by the incoming F_z and the Coriolis term at the bottom boundary, with the fluxes at the top boundary negligible in comparison. Noting the change in magnitude from the tropospheric box in Figure 5-7 to the stratospheric boxes in Figures 5-5 indicates that the wave flux entering the stratosphere represents a small loss to the troposphere but a large gain for the stratosphere.

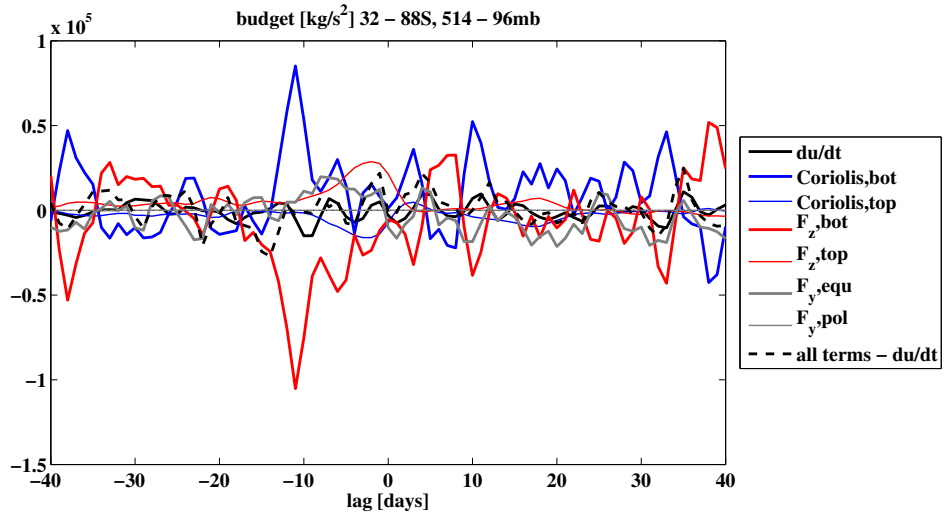

Figure 5-7: Same as Figure 5-5 but for the upper troposphere. The box is bounded at 514 hPa and 96 hPa at the top, and by lateral boundaries at 32°S and 88°S.

In an attempt to more clearly distinguish the relative contributions of the terms in the upper tropospheric momentum budget, the terms shown in Figure 5-7 are integrated in time over the duration of the composite (Figure 5-8). Again, a strong cancellation of the Coriolis term and the upward EP flux can be observed in the mid-troposphere. The terms at both the lower and the upper boundary do not exactly balance before the warming occurs, which is translated into a net momentum transport into the box from below as well as across the top boundary into the stratosphere. After the warming occurs, the terms balance more closely in magnitude, while they balance well in time throughout the composite. The change in the tropospheric winds is a small residual on the order of or smaller than the meridional EP flux. These terms emerge as a result of the imbalance in magnitude between the upward EP flux and

the Coriolis term. While the integrated budget does not reveal which terms cause the small change in the tropospheric momentum budget, it nevertheless gives a clearer picture of the balance of the terms in the upper troposphere.

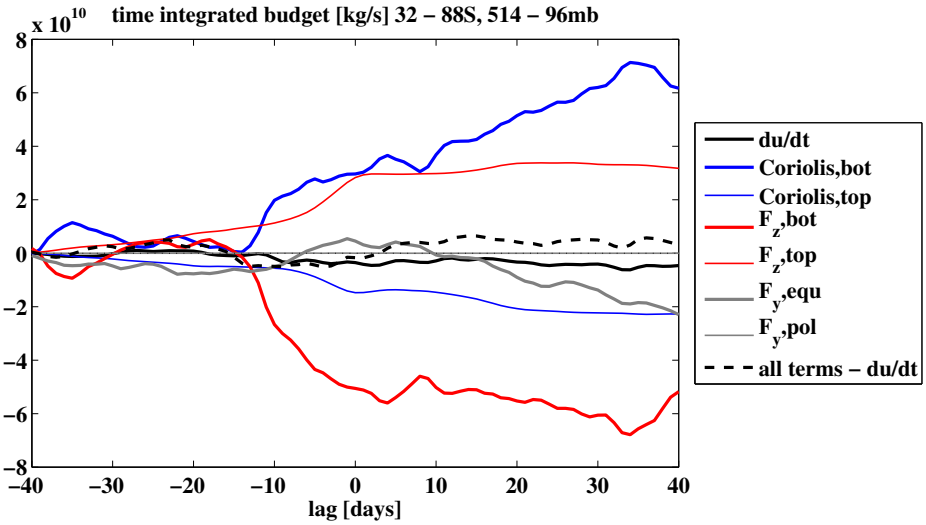

Figure 5-8: Same as Figure 5-7 but integrated in time starting on day -40.

Since a possible change in tropospheric momentum is still obscured by the large contributions from F_z and the Coriolis term, the temporal change in wind speed in the troposphere is considered separately in order to investigate in more detail the change to tropospheric relative momentum. The wind speed at the instantaneous (daily) tropospheric jet core increases by $1.2\,\mathrm{ms}^{-1}$ between day -9 to day -4, which corresponds to the period when the F_z fluxes in the lower stratosphere exhibit their fastest growth before the warming. However, this increase is small when comparing to the background variability of the tropospheric jet: The standard deviation of the maximum wind at the location of the instantaneous (daily) tropospheric jet maximum is $2.7\,\mathrm{ms}^{-1}$ over the entire model run, while for comparison, the standard deviation at the instantaneous location of the maximum zonal mean wind speed of the stratospheric vortex is $20.8\,\mathrm{ms}^{-1}$.

Concluding, the response to the loss of momentum to the stratosphere in terms of a change in strength of the tropospheric jet is very weak, partly due to the masking effect of the induced Coriolis force, but also due to the greater mass of the troposphere as compared to the stratosphere. It turns out, however, that part of the response to

the loss of momentum may trigger an annular mode response in terms of a change in the location of the jet core. Figure 5-1 shows a negative signature of the tropospheric annular modes for the couple of days before the criterion for a sudden warming is fulfilled in the stratosphere for both the upper troposphere and the surface.

In order to estimate this response, the integrated zonal mean zonal wind is considered separately in order to estimate the cumulative effect of the loss of easterly momentum in the troposphere. Based on the finding that the integrated EP flux determines the response in the stratosphere as opposed to the instantaneous flux (Scott and Polvani, 2004), one could expect a similar finding to hold for the change in zonal mean zonal wind change due to the integrated loss of easterly momentum. As shown in Figure 5-1, the troposphere experiences a negative annular mode response during and after the time of the warming, and the tropospheric negative annular mode response corresponds to an equatorward shift of the tropospheric jet (Figure 5-3). While the residual of the zonal mean zonal wind shown in figures 5-7 and 5-8 is too small to indicate a response in comparison with the wave fluxes, focussing on the integrated change in wind gives a clearer picture. Figure 5-9 shows the time integrated temporal change in zonal mean zonal wind (i.e. Equation (5.4)) integrated over the polar troposphere (bounded by 514 hPa at the bottom and 96 hPa at the top, and at 32° and 88°S) and the equatorial troposphere (514 to 96 hPa and 21° to 32°S). Equatorward of the jet maximum, the wind increases between the onset of the strong upward EP fluxes around about 10 days before until 20 days after the warming. Over the same time range, but slightly less continuous, the box poleward of the tropospheric jet shows a decrease in wind speed. These two effects indicate the onset of the annular mode response during the time of strong negative anomalies in the stratospheric annular mode index. After day 20, the wind response does not exhibit much change, indicating that the annular mode index stays in the observed negative pattern.

This compares to the finding of anomalous poleward F_y during the growth and mature phase of the warming, corresponding to equatorward momentum fluxes into the region of the tropospheric jet (Limpasuvan et al., 2004). These anomalous fluxes

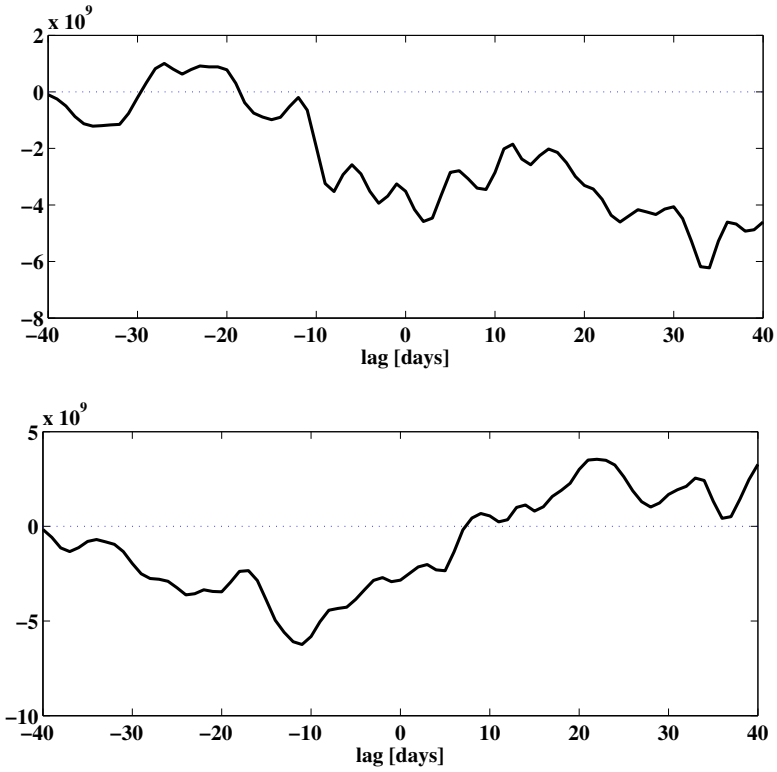

Figure 5-9: Time-integrated change in zonal mean zonal wind from Equation (5.4) integrated over the polar troposphere (top) and the equatorial troposphere (bottom). Units: $\mathrm{kg\,s^{-1}}$.

may be able to strengthen the tropospheric jet equatorward of its climatological latitudinal position.

Surface Response to Sudden Warmings in the Momentum Budget

In addition to the upper tropospheric response, the surface response in the sudden warming composite was investigated. The terms contributing to a box bounded at the lowest model levels are Rayleigh damping and topographic stress as described in Section 5.3.2 which generally act to decelerate the wind at the lower boundary. These terms are balanced by the Coriolis fluxes and vertical EP fluxes at the top of the box, as well as the meridional EP fluxes at the equatorward boundary.

Figure 5-10 shows the budget including the above terms. Note that these terms are not anomalies with respect to the longterm mean, but absolute terms in order to better illustrate the sign and magnitude of the terms, since they are of similar order. Rayleigh damping (green dashed line) acts to damp the wind in the box (i.e. it is a negative contribution to the box), while the topographic stress (gree solid line) depends strongly on the surface wind at the topography. At the top of the box, again the Coriolis term and the vertical EP flux balance very well in time and magnitude, with a net upward exchange of momentum. In general, the wind follows the topographic surface stress, however the magnitude of the topographic stress turns out to be too large to balance the fluxes on the upper and equatorward boundaries of the box.

Due to the imbalance of the terms in the surface budget, the evaluation of the budget was limited to the locations on top of the mountains for this chapter.

Comparison to the Truncated Model Run

The above analysis shows that the upward EP flux is well balanced by the contribution of the Coriolis term to the residual circulation in the troposphere at all times. While the above results suggest the setup of a tropospheric eddy feedback, in order to conclusively distinguish between an induced eddy feedback (Chen and Robinson, 1992) and the possibility for a meridional circulation reaching all the way to the sur-

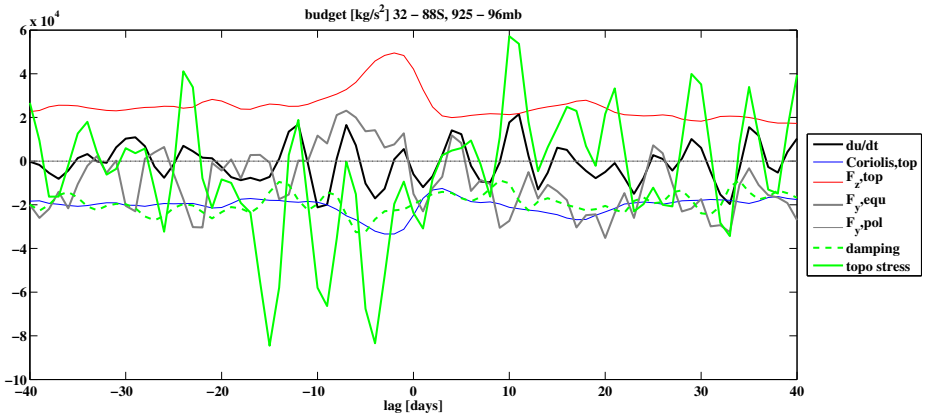

Figure 5-10: Same as Figure 5-7 but for the entire troposphere and in absolute numbers (not anomalies). The box is bounded at 925 hPa (the lowest model level) and 96 hPa at the top, and by lateral boundaries at 32°S and 88°S.

face as suggested by Thompson et al. (2006), it would be ideal to be able to compare to a model run without synoptic eddies in order to analyze the tropospheric response in the absence of the synoptic eddy feedback. The truncated model run described in Chapter 2 seems to be an ideal candidate: The truncated model run does not include synoptic eddies, it does not include a mountain, so one would not have to include form stress at the top of the mountain, and similar to the topography run, the signal from the sudden warming is observed throughout the stratosphere.

However, despite the absence of the synoptic eddies it turns out that a weak annular mode response is present in the troposphere for the truncated run, obscuring the effect of the momentum loss on the mean flow. This effect is likely due to the observed baroclinic instability of the planetary wave numbers described in Chapter 4 inducing a weak annular mode feedback in the troposphere. Figure 5-11 depicts the annular mode response of the truncated run for a composite of sudden warmings identified according to the PC3 criterion described in Section 3.3. While the annular mode response to the stratospheric signal is considerably smaller than for the topography run, the presence of a weak annular mode feedback does not allow for a more definitive conclusion. The corresponding EOF patterns (Figure 5-12) indicate the same variability observed in the topography run as indicated in Figure 2.2.3. Comparing to Figure 2-6 it can be seen that the node of the tropospheric EOF pattern (bottom

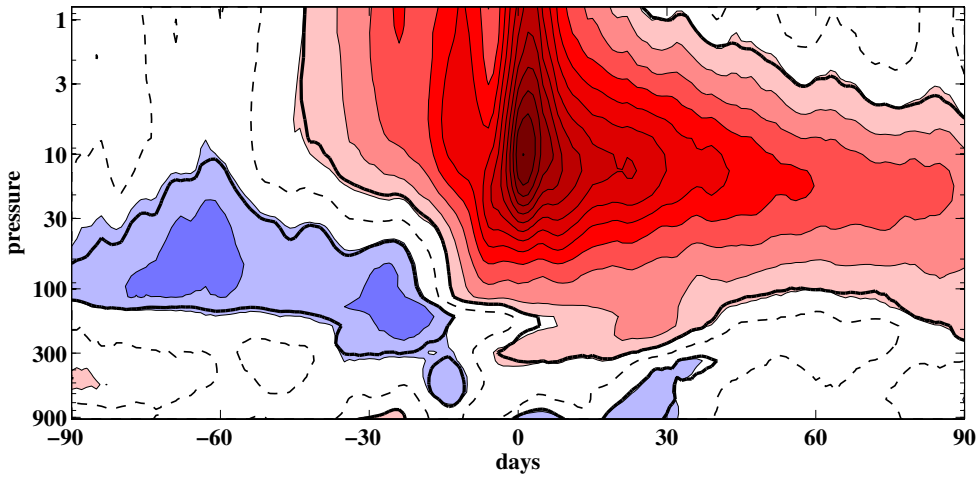

Figure 5-11: Same as Figure 5-1 but for the run truncated to wave-2 and a zonal mean flow as described in Section 2.2.3.

panel of Figure 5-12) lies on the tropospheric jet, indicating a meridional shift of the jet as the tropospheric mode, while the stratospheric pattern exhibits a maximum at the location of the stratospheric polar vortex, indicating a weakening / strengthening pattern as observed in the topography run.

These findings indicate that the tropospheric variability of the truncated run is similar to the variability in the topography run, yielding the truncated run unsuitable to explore tropospheric variability in the absence of a synoptic eddy feedback. As Chapter 4 shows, part of the role of synoptic baroclinic eddies is taken on by planetary-scale waves, which probably yields the annular mode - like response observed in the truncated run.

5.5 Discussion

In summary, it is observed that the EP fluxes and the residual circulation balance very well throughout the extratropical atmosphere down to the mid-troposphere during times of sudden warmings. A deposition of momentum induces a meridional circulation which on daily time scales offsets the effect of the wave flux to a major extent. In the stratosphere, the small imbalance between the upward EP flux and the Coriolis term yields the observed deceleration of the zonal mean wind during

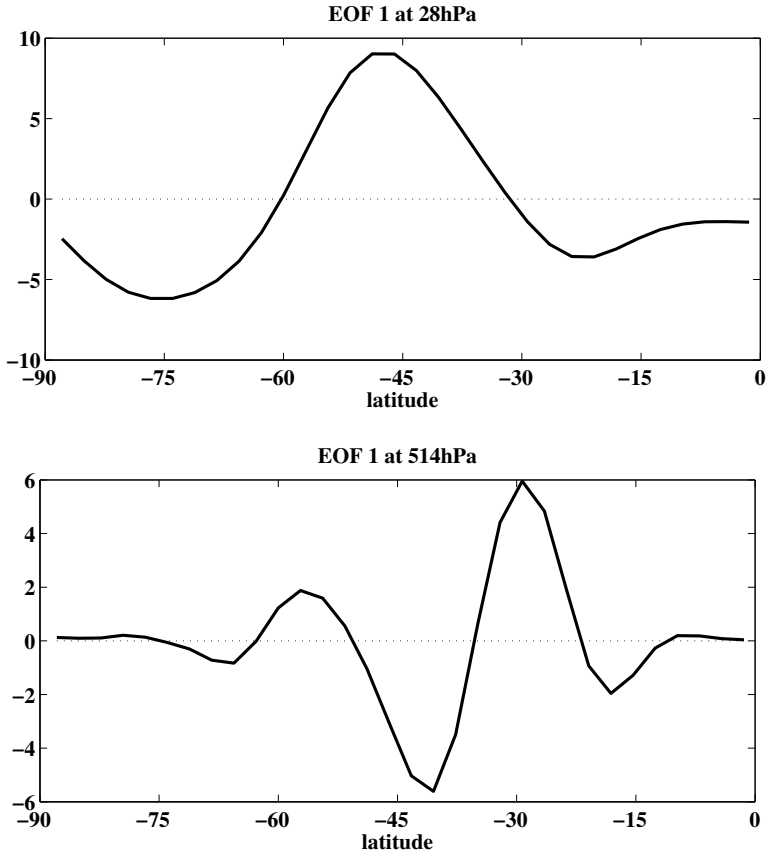

Figure 5-12: Same as Figure 5-3 but for the run truncated to wave-2 and a zonal mean flow as described in Section 2.2.3.

stratospheric sudden warmings. These results confirm calculations by Haynes et al. (1991), indicating an instantaneous response by the meridional circulation at the location of the forcing, while the downward extension of the meridional circulation cell happens on slower time scales and is dependent on the radiative relaxation time. The instantaneous response can be compared to the *Eliassen* response (see e.g. Haynes et al. (1991)), where the time scale of the forcing is much smaller as compared to the radiative time scale, as observed in the present model run. The Eliassen response can be expected to be valid in the case of sudden warmings considering the small time scales on the order of 10 days which is observed in this analysis as the time scale of the increased forcing by the waves, while also the predicted remote effect in the tropics can be observed in this budget.

In the troposphere, the close balance between the EP fluxes and the residual circulation almost entirely cancels the expected effect on the mean flow, while a strengthening of the zonal wind due to the loss of easterly momentum could be expected. The tropospheric change in zonal mean wind is very weak, as zonal wind turns out to be only a small residual in the balance between F_z and the residual circulation induced in the troposphere. The loss of tropospheric momentum over the 10 days before the stratospheric warming, induces a weak strengthening of the tropospheric jet on its equatorward flank and a small weakening of the jet at its poleward flank, inducing an annular mode response which is generally observed as a response to stratospheric forcings in the atmosphere, while here a weak response is induced by the loss of momentum to the stratosphere already. The annular mode index then remains negative (by convention, a weakening of the polar vortex in the stratosphere as well as an equatorward shift of the tropospheric jet correspond to a negative annular mode index) for an extended period of time up to two months after the warming has occurred in the stratosphere.

The lasting nature of the negative annular mode index in the troposphere indicates an exchange of information between the troposphere and the lower stratosphere, which exhibits long radiative time scales and which therefore recovers very slowly from the impact of the upper stratospheric signal. However, due to the close balance

of the wave flux and the residual circulation it is difficult to observe the effect of the stratospheric signal in the troposphere, which seems to be able to reinforce the negative annular mode pattern in the troposphere. These results confirm results by Song and Robinson (2004), who find that the stratospheric forcing is maintained by the interactions with tropospheric eddies, and both the induced residual circulation as well as the waves are equally responsible for transferring the signal between the stratosphere and the troposphere, making it impossible to talk about an induced residual circulation without looking at the waves as well.

It can therefore be concluded that the signal from the sudden warming in the upper stratosphere propagates into the lower stratosphere by a combination of the above mentioned mechanisms (wave reflection, downward control, or critical line interaction), impacting the lower stratosphere. An eddy feedback in the troposphere can then be observed in terms of the annular mode index signal, consistent with e.g. Song and Robinson (2004), Gerber and Polvani (2009), Polvani and Kushner (2002). This annular mode signal is likely caused by synoptic eddy feedback, which is indicated by the response in the meridional EP flux. On the other hand, the direct effect on the tropospheric wind is small. This indicates that a tropospheric eddy feedback is necessary for a tropospheric response to sudden warmings.

These conclusions indicate that the downward influence on the troposphere as described by Thompson et al. (2006) for the effect of ozone depletion is not the dominant effect present in the model atmosphere studied here, since the eddy feedback in the troposphere turns out to be an important part of the angular momentum balance. While the induced meridional circulation provides the major contribution towards balancing the upward wave flux, it does not directly affect the zonal mean wind speed. However, it is here not possible to evaluate the role of the induced meridional circulation for the longer forcing time scales studied in Thompson et al. (2006).

In summary, the net effect on the troposphere is found to be small, and the momentum budget is not able to identify a causal relationship between the change in wind speed and the remaining terms, although the forcing is transient. Therefore,

the mechanism of stratospheric influence on the troposphere after a stratospheric warming remains unclear due to the difficulty to distinguish between passive [Song and Robinson (2004), Charney and Drazin (1961)] and active (Chen and Robinson, 1992) mechanisms, since upward wave propagation is (*passively*) inhibited after the warming (i.e. the F_z flux into the stratosphere decreases, compare e.g. Figure 5-5), while in addition the upward wave flux into the stratosphere is (*actively*) balanced by the meridional circulation on a daily basis.

This study indicates that while momentum budget analysis gives a good indication of the balance of the respective terms in the momentum budget, future work will have to focus on the response of tropospheric synoptic eddies to stratospheric forcing and their indirect effect on tropospheric wind speeds.

Chapter 6

Principal Oscillation Patterns Applied to Stratosphere - Troposphere Coupling

6.1 Introduction

In order to further investigate the coupling between the stratosphere and the troposphere, Principal Oscillation Pattern Analysis is introduced as a tool to capture the dominant modes of variability of the stratosphere - troposphere system. The goal of the analysis is to ultimately be able to predict the response of a forcing onto the atmosphere depending on its location and strength. Such a forcing could e.g. be provided by an Eliassen-Palm flux convergence in the upper stratosphere as observed during sudden warming events. While it is desirable to be able to estimate the magnitude and location of the response given the magnitude and location of the forcing, the previous chapters have shown that the atmospheric response to a stratospheric sudden warming is not straight forward even when evaluating every term in a momentum budget framework. Since the exact processes and mechanisms of coupling and momentum exchange are not resolved, the problem of stratosphere - troposphere coupling is here approached from a statistical viewpoint.

In order to understand how a stratospheric forcing projects onto the troposphere, a first step has to investigate the internal variability of the stratosphere and the troposphere, as well as a possible co-variability between the respective layers.

The variability inherent to the extratropical atmosphere for both observational as well as modeling studies is often framed using Empirical Orthogonal Function (EOF) analysis to detect the dominant patterns of variability. EOF patterns are computed by decomposing a data set into orthogonal basis functions, defined as the eigenvectors of the covariance matrix of the data. Supposing that these eigenvectors depict the dominant variability of the atmosphere, an atmospheric forcing (e.g. a heat or a momentum forcing) is expected to project onto the EOF patterns.

However, while EOF analysis is a valuable tool to detect dominant patterns in a dataset, these patterns do not account for temporal changes in the data. Lagged regression analysis can be used as a tool to find time-dependent structures, and is successful at detecting correlations between stratospheric and tropospheric patterns at different lags, however detecting a time-dependent structure in the first place depends on *guessing* the correct lag. Moreover, EOF patterns are highly sensitive to normalization and weighting (North et al., 1982), and in some cases, robust patterns can not be obtained [e.g. Taguchi (2003)]. This finding suggests that EOF analysis is not consistently able to detect the modes and the dynamical time-dependent features of a system.

In a second step, it is desirable to quantify the response to a forcing onto the atmosphere. However, the above indication that EOF patterns do not reliably detect the modes of a system questions their ability to quantify the response to an external forcing. If the effect of a forcing and the time scales of variability are to be quantified, an improved description of the inherent dynamics of the system is required.

This study suggests that Principal Oscillation Patterns (POP) provide an analysis yielding the intrinsic variability of the system and a second pattern corresponding to the mode which the forcing tends to project onto. These patterns are the eigenmodes of a dynamical operator describing the evolution of the flow and involving the dynamical mechanism. POP analysis applied to extratropical atmospheric variability

is expected to detect the dynamical modes and the time-dependent structure of the system. It is anticipated that when the projection of a forcing onto the atmosphere is not described by an EOF pattern, then the modes provided by POP analysis will differ from the EOF patterns.

We suggest that POP analysis as introduced in the next section is able to overcome the shortcomings of EOF analysis, while in addition providing a quantitative and dynamical way of characterizing the eigenmodes of the atmosphere as well as the response to an external forcing. The following sections will explain the applicability of POP analysis to zonal mean zonal wind variability in the extratropical atmosphere. The analysis of Ring and Plumb (2008) is extended to a statistical analysis of zonal mean zonal wind in the stratosphere, hoping to better relate the forcing and response between the troposphere and the stratosphere.

This chapter will provide the findings of POP analysis for the unforced atmosphere and therefore for internal atmospheric variability, and thereby build the basis for work with forced model runs. The forced runs will be studied in future work.

This chapter is organized as follows: The next section will introduce Fluctuation - Dissipation analysis, Section 6.2 will introduce the analysis performed in this study and propose Principal Oscillation Pattern analysis as a statistical tool to quantify the internal variability of the analyzed data. The model data will be introduced in Section 6.3, and a test of the analysis' stability and robustness will be performed and documented in Section 6.4. The results of the analysis will be shown in Section 6.5, while Section 6.6 offers a discussion of the results.

6.2 Principal Oscillation Pattern Analysis

6.2.1 The Fluctuation - Dissipation Framework

As indicated in Chapter 5 and other model studies [e.g. Ring and Plumb (2007), Ring and Plumb (2008)], simple dynamical models show that an annular mode response can be triggered by applying external forcings onto the extratropical atmosphere. The

tropospheric response is communicated and organized by tropospheric eddy feedback (Lorenz and Hartmann, 2001). The dependence on the nature, location and strength of the forcing has been analyzed in previous studies without an explicit characterization of the mechanism using simplified atmospheric general circulation models [Ring and Plumb (2008), Ring (2008), Chan and Plumb (2009)]. A linear relationship between a heat or a momentum forcing and its respective response in zonal mean zonal wind was found [e.g. Ring and Plumb (2008), Ring and Plumb (2007), Gerber et al. (2008)]. These results justify the use of Fluctuation - Dissipation analysis, which linearly relates the forcing onto an atmospheric mode to its response through a characteristic time scale.

Fluctuation - Dissipation analysis was introduced to climate sciences by Leith (1975) and relates the internal variability of the system to a sufficiently small external forcing, assuming that a projection of the forcing onto the internal variability of the system scales linearly with the response of the system projected onto the intrinsic mode.

$$\text{Mode} \cdot \text{Response} = \tau \cdot (\text{Mode} \cdot \text{Forcing}). \tag{6.1}$$

The associated characteristic time scale τ gives the slope of this linear relationship for each mode. Equation (6.1) indicates that in order to be able to estimate the response to a given forcing, the modes of the system have to be explicitly defined. In the context of this analysis, the fluctuations about that mean state are given by stochastic stirring by baroclinic eddies (Vallis et al., 2004).

In Ring and Plumb (2008), a change in the forcing amplitude is shown to yield a linear increase in response as assumed by the Fluctuation - Dissipation theorem. The amplitude of the response could also be altered by a change in location of the forcing, changing the projection of the forcing on the mode (Ring and Plumb, 2007), which also yielded linear results.

6.2.2 Introduction to POP Analysis

Principal Oscillation Pattern analysis was first introduced to climate sciences by von Storch et al. (1988) and Hasselmann (1988). We will here follow the approach given in Penland (1989). POP analysis is currently used for e.g. seasonal prediction of tropical sea surface temperatures (Alexander et al., 2008).

A stochastically forced system following a linear relationship may be described by a first order autoregressive process as given by

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{B}\,\mathbf{x} + \mathbf{f}\,. \tag{6.2}$$

Bold lower case letters denote column vectors, bold upper case letters denote matrices. $\mathbf{B}$ represents a dynamical operator governing the propagation of the system and $\mathbf{x}$ is the state vector. The forcing term $\mathbf{f}$ is assumed to be uncorrelated to the system variable in $\mathbf{x}$ and to decorrelate much faster than $\mathbf{x}$. The forcing acts to maintain the variance of the system, while the internal dynamics act to dissipate it (Newman and Sardeshmukh, 2008). It is here approximated as a stochastic or random forcing in time and space for the entire domain.

Formally solving the stochastic differential Equation (6.2) yields

$$\mathbf{x}(t_0 + \tau) = \mathbf{G}(\tau)\,\mathbf{x}(t_0) + \mathbf{f}(t_0, \tau),$$

where τ is a lag after the initial time t_0, $\mathbf{G}$ is the propagation operator in time $\mathbf{G}(\tau) = e^{\mathbf{B}\tau}$ and

$$\mathbf{f}(t_0, \tau) = \mathbf{G}(\tau) \int_{t_0}^{t_0+\tau} \mathbf{G}(s)^{-1}\mathbf{f}(s)\, ds$$

(Penland, 1989). For zero or white noise, the best estimate for the state vector $\mathbf{x}$ at lag τ is

$$\mathbf{x}(t_0 + \tau) = e^{\mathbf{B}\tau}\,\mathbf{x}(t_0). \tag{6.3}$$

The state vector $\mathbf{x}$ at lag τ can be expressed as $\mathbf{x}(t_0+\tau) = \mathrm{Re}\left[\sum_j u^j\, e^{\lambda^j \tau}\right]$ with the j-th element

$$\mathrm{Re}\left[u^j\, e^{\lambda^j \tau}\right] = \mathrm{Re}\left[(u_r + iu_i)^j \cdot e^{(\lambda_r + i\lambda_i)^j \tau}\right] = \left[u_r^j\, \cos(\lambda_i^j \tau) - u_i^j\, \sin(\lambda_i^j \tau)\right] e^{\lambda_r^j \tau} \quad (6.4)$$

where u^j represents the j-th eigenvector of $\mathbf{B}$ and λ^j is the corresponding eigenvalue (Penland and Sardeshmukh, 1995). Subscripts r and i denote real and imaginary parts, respectively. For a purely real eigenvector u^j, Equation (6.4) reduces to $u_r^j\, e^{\lambda_r^j \tau}$, which represents a purely decaying structure for $\lambda_r < 0$. For a non-zero imaginary part of $\mathbf{x}$, two subsequent POP modes are complex conjugate and the mode corresponds to an oscillation between their real and the imaginary parts, with an evolution $u_r \to -u_i \to -u_r \to u_i \to u_r$, modulated by $e^{\lambda_r \tau}$. The system eigenvectors can therefore exhibit decaying as well as oscillating modes. In addition, propagation can be observed considering two subsequent modes, if the spatial patterns exhibit features in quadrature to each other, and if in addition the time series of these modes are in quadrature with respect to each other.

6.2.3 Advantages of POP Analysis

In order to obtain the state vector at lag τ from Equation (6.3), the dynamical operator $\mathbf{B}$ has to be determined. $\mathbf{B}$ can be decomposed using eigenvector decomposition: $\mathbf{B} = \mathbf{V}\,\mathbf{\Lambda}\,\mathbf{W}^\top$ with $\mathbf{\Lambda}$ a diagonal matrix of eigenvalues $\lambda_{1,\ldots,n}$ and $\mathbf{V}\mathbf{W}^\top = \mathbf{V}^\top\mathbf{W} = \mathbf{I}$, with $\mathbf{I}$ the identity matrix and $^\top$ the transpose complex conjugate. Then $\mathbf{V}$ contains the right eigenvectors and $\mathbf{W}^\top$ the left eigenvectors of the dynamical operator $\mathbf{B}$.

The strength of POP analysis is that the matrices $\mathbf{V}$ and $\mathbf{W}$ containing the eigenvectors of the system matrix can be obtained entirely from the data contained in the state vector $\mathbf{x}$. Starting from the lagged covariance of the system variable

$$\mathbf{C}_\tau = < \mathbf{x}(t_0+\tau)\, \mathbf{x}(t_0)^\top >$$

(with $< . >$ the time average) and mapping it from eigenspace according to $\mathbf{x} = \mathbf{V}\bar{\mathbf{x}}$ and from real space to eigenspace by using $\bar{\mathbf{x}} = \mathbf{W}^\top \mathbf{x}$ (variables with overbar are in eigenspace, all other variables are in physical space), this yields for the lagged covariance in eigenspace: $\bar{\mathbf{C}}_\tau =< \bar{\mathbf{x}}(t_0+\tau)\,\bar{\mathbf{x}}(t_0)^\top >$. Following Ring (2008) it can be shown that

$$\bar{\mathbf{C}}_\tau = \mathbf{\Gamma}\,\bar{\mathbf{C}}_\mathbf{0} \quad \text{for } \tau > 0 \tag{6.5}$$

where $\mathbf{\Gamma}$ is a diagonal matrix with elements $e^{\lambda_i \tau}$. Using (6.5) along with the described transformations yields for $\tau > 0$

$$\begin{aligned} \mathbf{C}_\tau &= \mathbf{V}\,\bar{\mathbf{C}}_\tau \\ &= \mathbf{V}\,\mathbf{\Gamma}\,\bar{\mathbf{C}}_\mathbf{0} \\ &= \mathbf{V}\,\mathbf{\Gamma}\,\mathbf{W}^\top\,\mathbf{C}_\mathbf{0} \\ &=: \mathbf{G}_\tau\,\mathbf{C}_\mathbf{0}\,, \end{aligned}$$

where $\mathbf{G}_\tau := \mathbf{V}\,\mathbf{\Gamma}\,\mathbf{W}^\top$, so that the matrices $\mathbf{V}$ and $\mathbf{W}$ are the right and left eigenvectors not only of $\mathbf{B}$, but also of $\mathbf{G}_\tau$, which can be computed directly from the lagged and the zero-lag covariance matrices using $\mathbf{G}_\tau = \mathbf{C}_\tau\,\mathbf{C}_\mathbf{0}^{-1}$. The eigenvalues of $\mathbf{G}_\tau$ are related to the eigenvalues of $\mathbf{B}$ through $\mathbf{\Gamma} = e^{\mathbf{\Lambda}\tau}$.

The POP eigenmatrix $\mathbf{V}$ describes the intrinsic variability of the system. This can be seen using the decomposition for $\mathbf{B}$ in Equation (6.2)

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{V}\mathbf{\Lambda}\mathbf{W}^\top\,\mathbf{x} + \mathbf{f}. \tag{6.6}$$

Considering the unforced problem ($\mathbf{f}$=0), using $\mathbf{W}^\top\mathbf{x} = \bar{\mathbf{x}}$ and multiplying (6.6) by $\mathbf{W}^\top$ from the left yields

$$\frac{\partial \bar{\mathbf{x}}}{\partial t} = \mathbf{\Lambda}\,\bar{\mathbf{x}}$$

with a solution $\bar{\mathbf{x}} \sim e^{\mathbf{\Lambda}t}$. With $\mathbf{x} = \mathbf{V}\bar{\mathbf{x}}$ this yields $\mathbf{x} \sim \sum_n \mathbf{v_n} e^{\lambda_n t}$, where $\mathbf{V}$ describes the internal variability of the system. On the other hand, the POP eigenmatrix $\mathbf{W}^\top$ describes the pattern which the forcing projects onto. This can be seen from Equation

(6.6) by multiplying with $\mathbf{W}^\top$ from the left which yields

$$\frac{\partial}{\partial t}(\mathbf{W}^\top \mathbf{x}) = \mathbf{\Lambda} \mathbf{W}^\top \mathbf{x} + \mathbf{W}^\top \mathbf{f}$$

where the forcing as well as the state vector project onto $\mathbf{W}^\top$.

A further advantage of POP analysis is that the characteristic time scale of a mode is given directly by the eigenvalue of the respective mode. This can be seen from Equation (6.4), where $-1/\lambda_r$ gives the decay time scale of a mode, which can directly be computed from the output of POP analysis.

6.2.4 Computing POP Modes from Zonal Wind

The state vector $\mathbf{x}$ can span the entire variable space considered relevant to the analysis (i.e. zonal and meridional wind, temperature and surface pressure), however following Ring and Plumb (2008) only zonal wind was considered here in order to keep the analysis clear and for an easier subsequent interpretation of the results. In order to reduce the variable space to zonal wind as the only variable in the state vector $\mathbf{x}$, additional approximations and balance assumptions have to be made as described in Ring and Plumb (2008): Starting from an angular momentum budget in the primitive equations they derive an equation for zonal wind by assuming gradient wind balance and by defining an *effective torque*, which is designed to include both the applied momentum forcing as well as the advection of angular momentum by the residual circulation (the *Eliassen response* discussed in Chapter 5), which broadly corresponds to a torque extending below the applied forcing given by EP flux convergence. The most restrictive assumption is writing the anomalies in the EP flux divergence as a linear function of the zonal mean wind anomalies. Ring and Plumb (2008) find that these assumptions are justified for a model run comparable to the present control run. The detailed derivation of the reduced state vector can be found in the appendix of Ring and Plumb (2008).

For this configuration, Equation (6.2) describes the temporal change of the zonal mean wind forced by baroclinic eddies represented by the forcing $\mathbf{f}$. The dynamic

operator $\mathbf{B}$ includes the linear dynamics in terms of the mean winds as well as an eddy feedback mechanism as described in Song and Robinson (2004) and Ring and Plumb (2008). As a parameterization for the eddy feedback is not available, $\mathbf{B}$ is left to be determined from the data $\mathbf{x}$. POP analysis provides a framework for determining $\mathbf{B}$ entirely from the time evolution of the state vector.

The zonal wind data from the control run (described in Chapter 2) is averaged in longitude and the time mean of the data is removed, yielding the wind anomalies as a function of latitude, pressure level and time: $\mathbf{U}'(\varphi, p, t)$ with a length of $t = t_{max}$, where φ is latitude, p is pressure and t is time. For numerical reasons, the anomalies are framed in a two-dimensional matrix using a position index $q = \varphi + \varphi_{max}(p-1)$. The symmetric zero-lag covariance matrix and the lagged covariance matrix are computed as

$$\mathbf{C_0}(q,q) = \frac{\mathbf{U}'(q,t)\,\mathbf{U}'(q,t)^\top}{t_{max}} \text{ and } \mathbf{C}_\tau(q,q) = \frac{\mathbf{U}'_\tau(q,t)\,\mathbf{U}'(q,t)^\top}{t_{max}}$$

where $\mathbf{U}'_\tau(q,t)$ is the lagged time series and $\mathbf{U}'(q,t)$ is the time series at zero lag. Lags τ of the order of typical extratropical decorrelation time scales are used. The asymmetric $\mathbf{G}_\tau$ matrix is computed as the product of the inverse zero-lag and the lagged covariance matrices

$$\mathbf{G}_\tau(q,q) = \mathbf{C}_\tau(q,q)\,(\mathbf{C_0}(q,q))^{-1}. \tag{6.7}$$

The POP eigenvectors are the left and right eigenvectors of $\mathbf{G}_\tau$

$$\mathbf{G}_\tau(q,q) = \mathbf{V}(q,n)\,\mathbf{\Gamma}(n,n)\,\mathbf{W}(q,n)^\top,$$

where n is an index indicating the mode number $n = 1, ..., q_{max}$. $\mathbf{V}$ and $\mathbf{W}^\top$ are bi-orthogonal, but a normalization has to be applied to ensure that $\mathbf{V}^\top\mathbf{W} = \mathbf{I}$, where $^\top$ denotes the transpose complex conjugate, as $\mathbf{V}$ and $\mathbf{W}^\top$ can have complex entries. Since $\mathbf{G}$ is real, the POP eigenvectors are either real eigenvectors with real eigenvalues or eigenvectors in complex conjugate pairs with associated complex conjugate pairs of eigenvalues. The corresponding principal component time series $\mathbf{PC}(n,t)$ for each

mode n can be computed from

$$\mathbf{PC}(n,t) = \mathbf{W}(q,n)^\top \mathbf{U}'(q,t), \tag{6.8}$$

yielding that the original data can be reconstructed from

$$\mathbf{U}'(q,t) = \mathbf{V}(q,n)\,\mathbf{PC}(n,t) \tag{6.9}$$

by multiplying Equation (6.8) with $\mathbf{V}$ from the left and using $\mathbf{V}\mathbf{W}^\top = \mathbf{I}$. The error covariance matrix $\mathbf{Q}$ can be computed following Penland (1989):

$$\mathbf{Q}(q,q) = -\mathbf{B}(q,q)\,\mathbf{C_0}(q,q) - \mathbf{C_0}(q,q)\,\mathbf{B}(q,q)^\top.$$

The matrices $\mathbf{V}$ and $\mathbf{W}^\top$ can then be transformed back to depend on latitude φ, pressure p and mode number n.

6.3 Model Data

The zonal mean zonal wind data from the control model run (see Chapter 2) is used as an input to POP analysis. This run includes no external forcing, allowing for an analysis of the intrinsic atmospheric variability. The analysis includes 32 pressure levels between 1.9 and 925 hPa (excluding the sponge layer in the top layers) and 25 latitude bands between 20 and 87°S, i.e. the winter hemisphere.

The data is weighted in latitude to account for the decreasing surface area going towards the pole, as will be further described in Section 6.5.4. An additional weighting with respect to pressure will be discussed in Section 6.5.4. While the latitudinal weighting of the raw data does not change the patterns in a significant way, it is important to note that any decision on how to weigh or not weigh the data may have important consequences on the resulting patterns. For this reason, two types of pressure weighting will be compared in Section 6.5.4: raw model data on the model sigma levels interpolated to pressure, and the data on the same levels but weighted

by the respective mass that a level accounts for.

6.4 Analysis of POP Performance

Several methods have been proposed to test the performance of POP analysis and to evaluate its sensitivity. Among them are the τ-*test* (Penland, 1989) as well as a sensitivity test for the inversion of the zero-lag covariance matrix (Zhang et al., 1997).

In the following section, POP analysis performance will be investigated using Penland's τ-test as a consistency check for the analysis, while the inversion of the covariance matrix in Equation (6.7) will be evaluated using singular vector decomposition to compute the pseudoinverse of the zero-lag covariance matrix.

6.4.1 Penland's τ-Test

The τ-*test* proposed by Penland (1989) and documented in e.g. Penland and Sardeshmukh (1995) employs a key result of POP analysis. According to Equation (6.3), the lagged state vector $\mathbf{x}(t_0 + \tau)$ can be computed from the state vector at zero lag using the propagation operator $e^{\mathbf{B}\tau}$. The propagation operator can be estimated from a chosen lag and be tested for other relevant lags. If a system successfully reproduces the predicted lagged state vector, the τ-test justifies further dynamical diagnosis of the system using POP analysis, while a failure of the τ-test indicates that the chosen state vector is incomplete or that POP analysis is inadequate for the considered system (Newman and Sardeshmukh, 2008).

Although POP analysis is designed to be independent of lag, forecasting according to Equation (6.3) exhibits an increasing error for growing lag. This discrepancy can be predicted yielding the *theoretical discrepancy*:

$$\delta_{\text{theor}}(\tau) := \frac{\text{tr}(\mathbf{C_0} - \mathbf{G}_\tau \, \mathbf{C_0} \, \mathbf{G}_\tau^\top)}{\text{tr}(\mathbf{C_0})} \tag{6.10}$$

(Penland, 1989), where tr is the trace. This function is designed to increase from zero at $\tau = 0$ to 1 as τ goes to infinity. This theoretically predicted error can now

be compared to the *estimated discrepancy* which is computed directly from the data using the prediction of the propagation operator:

$$\delta_{\text{est}}(\tau) := \frac{< (\mathbf{x}(t_0+\tau) - \mathbf{G}_\tau \mathbf{x}(t_0))^\top (\mathbf{x}(t_0+\tau) - \mathbf{G}_\tau \mathbf{x}(t_0)) >}{\text{tr}(\mathbf{C_0})} \tag{6.11}$$

(Penland, 1989). A comparison of the theoretical and estimated discrepancies is

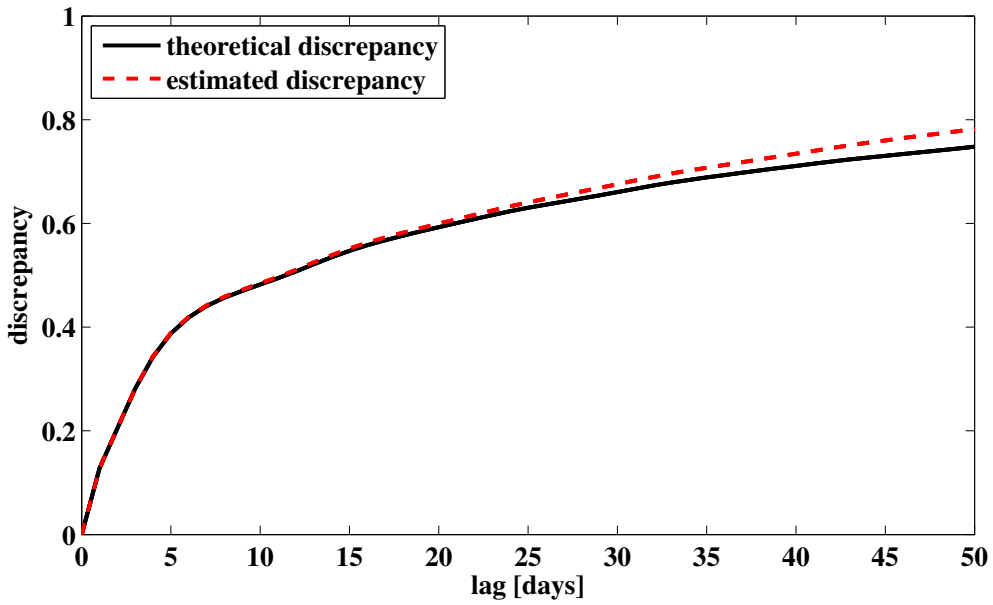

Figure 6-1: Theoretically predicted error (solid black curve) from Equation (6.10) and estimated error (dashed red curve) from Equation (6.11) for lags up to 50 days. A lag of 40 days will be used in the present analysis.

shown in Figure 6-1 for lags up to 50 days. Although the relative difference between the theoretical and the estimated error is shown to increase with lag, the relative error is sufficiently small, on the order of 3% for the lag of 40 days which will be employed for the present analysis (see next section). We therefore conclude that according to the τ-test, POP analysis is able to yield consistent results for the data used here.

6.4.2 Testing the Sensitivity to the Covariance Matrix Inversion

As the pressure levels and latitude grid points in the state vector $\mathbf{x}$ are not independent, inverting the zero-lag covariance matrix $\mathbf{C_0}$ as required by POP analysis in Equation (6.7) can lead to numerical difficulties as $\mathbf{C_0}$ does not exhibit full rank up to numerical precision. This numerical problem is usually dealt with by projecting the data onto a number of dominant EOFs before POP analysis is applied (Penland, 1989). However, the number of EOFs to retain remains arbitrary, and it was found that using EOF truncation for Fluctuation - Dissipation analysis on zonal wind data introduces a significant bias (Cooper and Haynes, 2008). In addition, for the subsequent comparison of the performance and patterns to EOF modes it is not desirable to bias the POP patterns using EOF truncation. Instead, a *pseudoinverse* of the matrix is computed, a method to not only test the sensitivity of the inversion as shown in Zhang et al. (1997), but to overcome the numerical difficulties.

The zero-lag covariance matrix $\mathbf{C_0}$ can be decomposed using singular vector decomposition $\mathbf{C_0} = \mathbf{MDM}^\top$. The diagonal of $\mathbf{D}$ consists of the eigenvalues $d_1, ..., d_n$ of the matrix $\mathbf{C_0}$ and $\mathbf{M}^\top\mathbf{M} = \mathbf{I}$, while the inverse is given by ${\mathbf{C_0}}^{-1} = \mathbf{MD}^{-1}\mathbf{M}^\top$. Assuming all eigenvalues of index greater than k are zero to numerical precision, using only the non-zero eigenvalues $d_1, ..., d_k$ of $\mathbf{C_0}$ for computing the inverse does not change ${\mathbf{C_0}}^{-1}$:

$$ {\mathbf{C_0}}^{-1} = \mathbf{MD}^{-1}\mathbf{M}^\top = \mathbf{MD}^{-\mathbf{1}}_{\mathbf{1},..,\mathbf{k}}\mathbf{M}^\top = \tilde{\mathbf{C}}_{\mathbf{0}}^{-1}. $$

However, if the truncated eigenvalues $d_{k+1}, ..., d_n$ are small in comparison to the eigenvalues $d_1, ..., d_k$, but not zero, the above relationship does not hold exactly and it has to be determined which amount of variance has to be retained (i.e. what value of k to use) in order to overcome the difficulties with the numerical inversion. An unambiguous truncation value k often cannot be determined, as noise and data cannot be separated exactly. Therefore, there is a range of possible values for k which yield stable solutions, and the choice of an optimal k depends on the pattern deemed most useful for the analysis, which remains arbitrary to a certain degree (Wunsch, 2006).

It was found for this analysis that a robust value of variance kept is 87%, as will be discussed in Section 6.5.2.

6.5 Results

This section describes the results obtained in this study in terms of the relevant POP patterns. A comparison of EOF and POP eigenvectors is provided as well as a sensitivity analysis.

6.5.1 POP Modes from Model Data

Figures 6-2 and 6-3 show the real and imaginary parts of the $\mathbf{V}$ and $\mathbf{W}^\top$ matrices for the first two POP modes (filled contours). The bold line separates positive from negative patterns, while the sign of the patterns is arbitrary, but consistent between $\mathbf{V}$ and the corresponding $\mathbf{W}^\top$ matrix and between the POP pattern and its corresponding $\mathbf{PC}$ time series. The white contours indicate the climatological zonal mean zonal wind. The $\mathbf{V}$ modes in Figure 6-2 show a dominantly stratospheric pattern, describing a latitudinal wobbling of the polar vortex around its climatological mean position. In the troposphere, the dominant pattern is a latitudinal wobbling of the tropospheric jet around its climatological mean position, but with a much weaker signature. According to these patterns, a poleward shift of the polar vortex is accompanied by an equatorward shift of the tropospheric jet and vice versa. The $\mathbf{W}^\top$ modes show a similar variability, but with a dominant tropospheric pattern. The imaginary parts of these patterns (Figure 6-3) exhibit a similar stratospheric variability, but the opposite connection to the troposphere. The remainder of this chapter will focus on the $\mathbf{V}$ patterns, since the $\mathbf{W}^\top$ patterns turned out to be not as robust to lag and retained variance. This will have an impact on the ability of the patterns to provide the modes for a forcing to project on, as will be discussed in the last section of this chapter. However, since only the internal variability of the model data is considered and analyzed here, this limitation will not affect the results obtained by the $\mathbf{V}$ patterns.

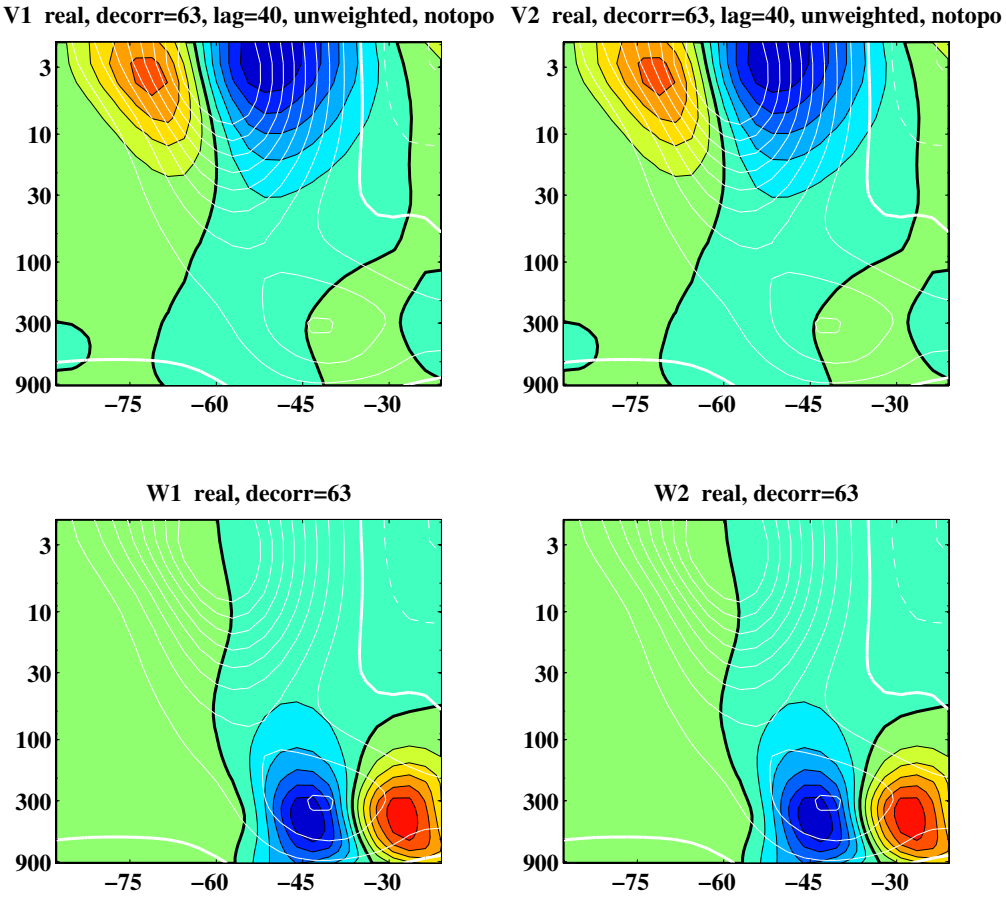

Figure 6-2: Filled contours: Real part of the POP patterns $\mathbf{V}$ (top) and $\mathbf{W}^\top$ (bottom) for the first (left) and second (right) mode. White contours: Zonal mean zonal wind averaged over the entire run, with a contour interval of $10\,\mathrm{ms}^{-1}$. The zero-wind line is bolded, negative contours are dashed.

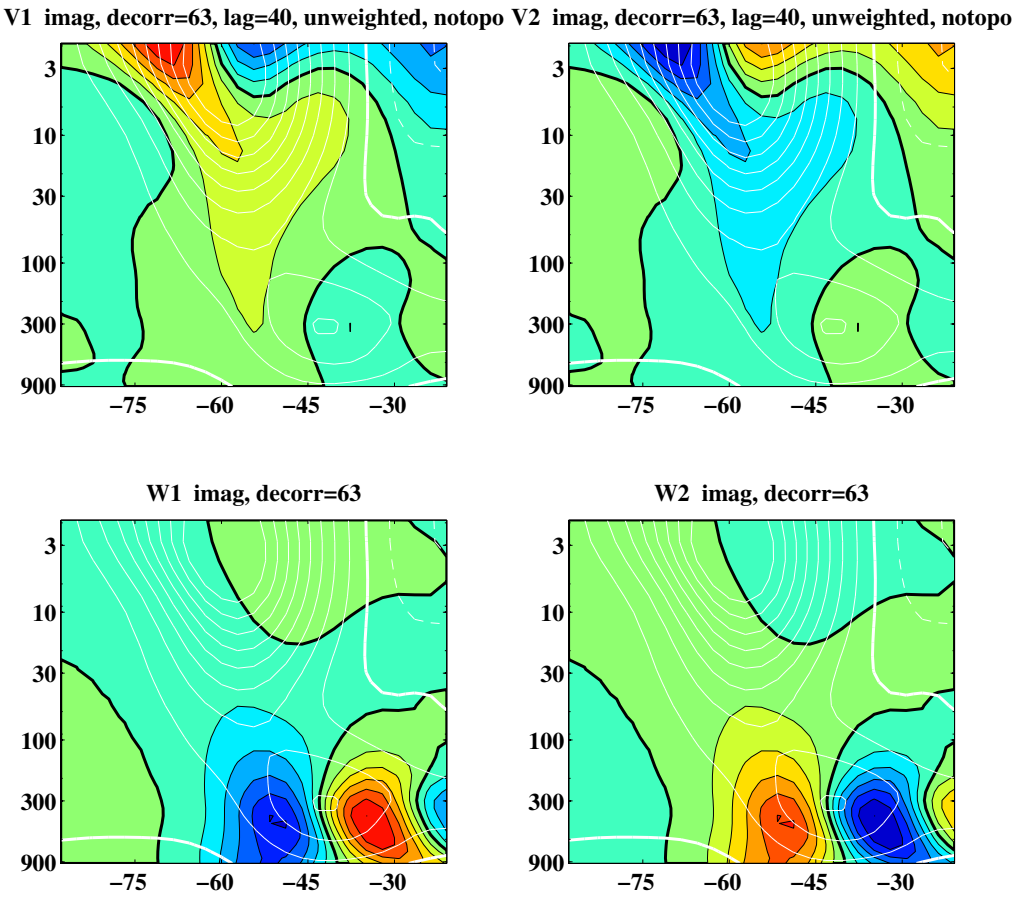

Figure 6-3: Same as Figure 6-2 but for the imaginary parts.

Since the patterns are complex, they exhibit an oscillating structure between the real and the imaginary patterns. Verifying the response in the oscillation of the real and imaginary parts of the principal component time series, it can be shown that the combination at times exhibits an oscillatory pattern with periods on the order of a hundred days, i.e. in the general form of $\mathbf{V}(q,n)\ \cos(t)+\mathbf{V}(q,n)\ \sin(t)$, corresponding to a progression as described in Section 6.2.2. At other times the time series stays in one phase for a while and intermittently oscillates back and forth between two dominant patterns. The phases of oscillatory motion correspond to times of stronger wind variability in the wind time series. Figure 6-4 shows an example of an oscillatory pattern corresponding to a 90 day excerpt of the control run. All observed oscillatory patterns exhibit the same direction of rotation.

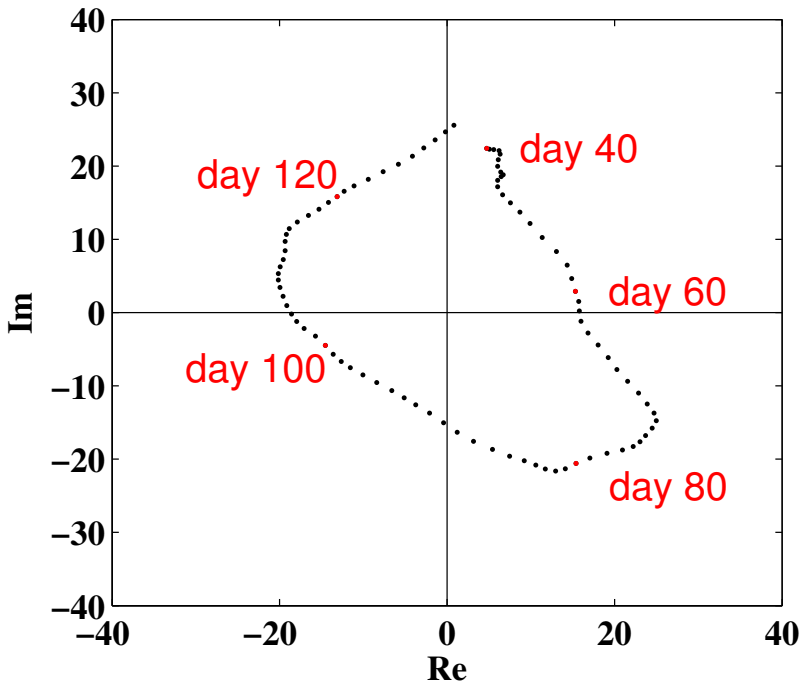

Figure 6-4: Days 40 to 130 of the principal component time series in the complex plane. Every dot represents the daily value of the principal component time series, with the annotated dots represented in red. The circular shape of the time series in the complex plane is indicative of an oscillatory behavior of the modes.

More specifically, the complex patterns can be reconstructed to the full dataset using Equation (6.9). Since the modes are complex, the POP modes as well as the time

series can be separated into real and imaginary parts (v_r, v_i) and (p_r, p_i), respectively, using $\mathbf{V}(q,1) = v_r + iv_i, \mathbf{V}(q,2) = v_r - iv_i$, and $\mathbf{PC}(1,t) = p_r + ip_i, \mathbf{PC}(2,t) = p_r - ip_i$. For including the first and second mode only, this yields

$$\begin{aligned}
\mathbf{U}'(q,t) &= \mathbf{V}(q,n)\,\mathbf{PC}(n,t) \text{ where } n = 1,2 \\
&= \mathbf{V}(q,1)\,\mathbf{PC}(1,t) + \mathbf{V}(q,2)\,\mathbf{PC}(2,t) \\
&= (v_r + iv_i)(p_r + ip_i) + (v_r - iv_i)(p_r - ip_i) \\
&= 2v_r p_r - 2v_i p_i. \qquad (6.12)
\end{aligned}$$

The reconstruction of the data from Equation (6.12) is shown in Figure 6-5 for the same time period as Figure 6-4. The data weighted by pressure (see Section 6.5.4) turns out to be more revealing for the connection to the tropospheric structure than the unweighted data, which is why the weighted pattern is shown here, while the unweighted pattern shows the same stratospheric pattern and the same tropospheric connection, but due to the weaker tropospheric magnitude the connection to the troposphere is not as obvious there. The progression of the patterns shows a strong emphasis on the mode describing the poleward shift of the polar vortex along with an equatorward shift of the tropospheric jet. Between days 40 to 60, the stratospheric pattern weakens along with an equatorward shift of the tropospheric pattern, leading to a reversal of the tropospheric jet shift. This reversal precedes the subsequent shift in the stratospheric pattern by a couple of days, yielding a pattern emphasizing an equatorward shift of the polar vortex along with a poleward shift of the tropospheric jet. Overall, the pattern corresponding to the poleward vortex shift and the equatorward jet shift dominates the time series of the oscillatory pattern, indicating that this pattern may be present most of the time, with a faster run through the opposite pattern. The analysis of other episodes of oscillatory motion between the first two POP modes indicates very similar results, with the longest-lived and dominant pattern being the one of a poleward vortex shift and the equatorward jet shift. This type of variability strongly resembles annular mode variability which is the dominant

pattern in the extratropical atmosphere (see Section 1.4.2). It is therefore not surprising that this pattern is found using POP analysis, it is however illuminating to see the progression of the pattern along with the corresponding time scales.

POP Modes from Forced Model Data

The modes for the control run showed a distinct latitudinal wobbling of the polar vortex along with a corresponding latitudinal movement of the tropospheric jet. This does not correspond to the annular mode pattern described in Section 1.4.2, which describes the dominant wintertime extratropical variability as a weakening/strengthening pattern of the stratospheric vortex along with a latitudinal wobbling of the tropospheric jet. Since stratospheric variability is considerably smaller in the control run as compared to the real atmosphere, as shown in 4, the dominant variability in the control run is not the characteristic weakening/strengthening pattern, but a latitudinal wobbling. This is an important finding for defining the internal variability of the stratosphere - troposphere system for weak tropospheric wave forcing.

In order to verify that the POP patterns pick up the variability which is characteristic to the winter stratosphere, POP and EOF patterns were also computed for the topography run described in Section 2.2.2 (not shown). Both the POP and EOF patterns for the topography run pick up the characteristic winter variability corresponding to a weakening (strengthening) of the stratospheric vortex along with an equatorward (poleward) shift of the tropospheric jet (not shown). This pattern corresponds to the mode observed in the topography run as shown in Section 5.4.1.

In order to correctly compute the projection of the forcing onto the POP modes for the forced system, an effective torque will have to be computed as described in Ring and Plumb (2008). This will be a matter of future research for this study.

6.5.2 Robustness to Variation in Lag and Retained Variance

The patterns in Figure 6-2 are sufficiently robust to variations in lag τ and variance kept for the computation of the pseudoinverse. This can be verified by checking

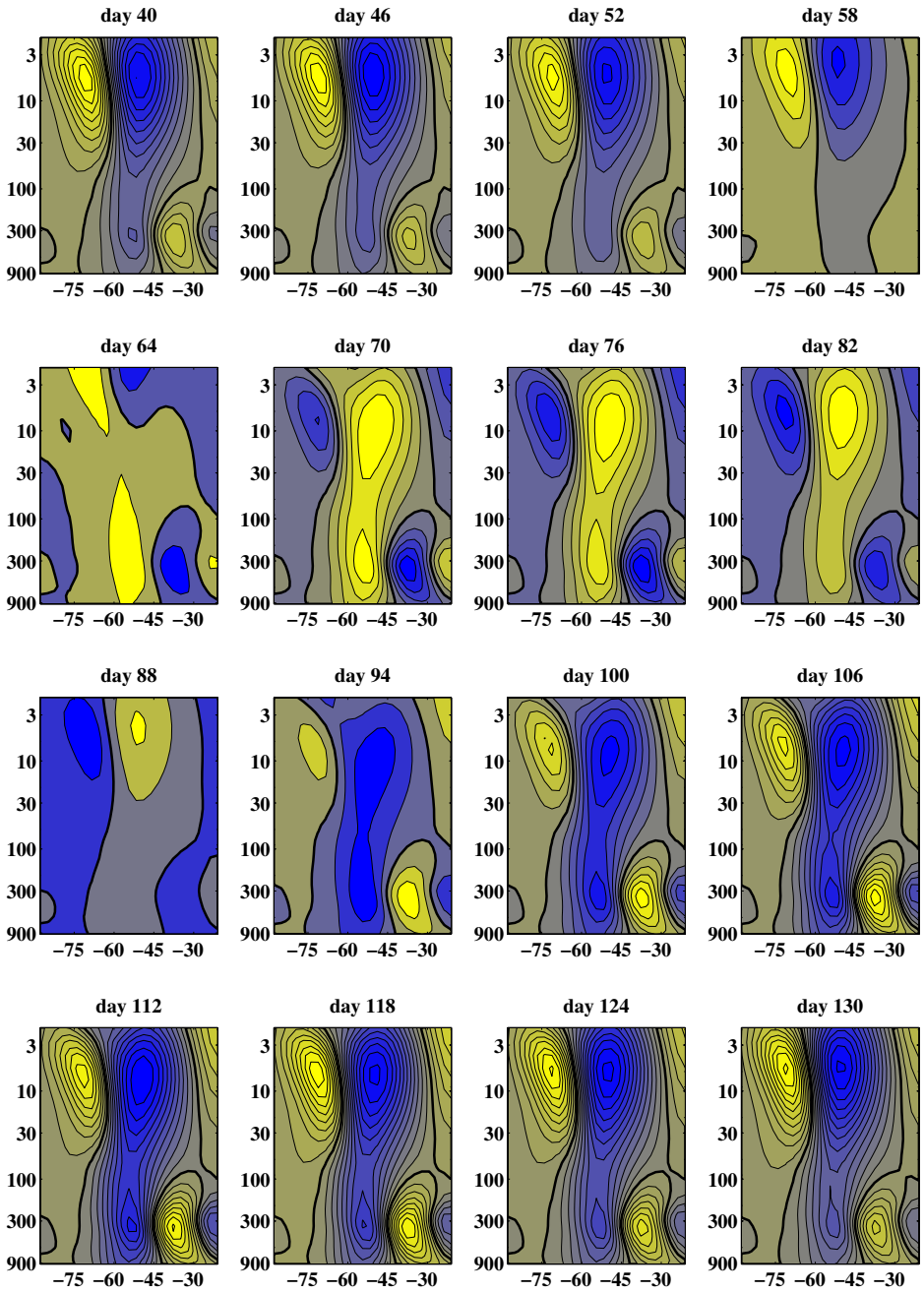

Figure 6-5: Progression of the oscillation between the real and the imaginary parts of the POP modes as given by Equation (6.12) over the same time range as Figure 6-4 (days 40 to 130), showing the weighted patterns. The yellow contours indicate positive values, blue contours indicate negative values. The contour interval is identical for all plots, indicating the relative difference in strength between the patterns.

Table 6.1: Comparison of characteristic time scales τ_{eig} computed from the eigenvalues of the POP analysis for varying lag and variance kept for the pseudoinverse. The time scale for the chosen variance and lag is highlighted.

	lag 20	**lag 40**	**lag 50**
var 84%	61	67	70
var 87%	58	63	67
var 91%	66	79	85
var 95%	66	79	85

the characteristic time scales of the modes using a range of lags and variance kept, while in addition comparing the characteristic time scales computed from the POP eigenvalues to the characteristic time scales from the corresponding time series. The intrinsic characteristic time scale of every mode is given by the real part of the POP eigenvalues in Equation (6.4), these time scales will be denoted τ_{eig}. The characteristic time scales given by the autocorrelation time scale of the corresponding principal component time series **PC** from Equation (6.8) will be denoted τ_{PC}.

Since the first two POP modes are complex conjugate, they exhibit the same characteristic time scales. The characteristic time scale τ_{eig} is 63 days for these modes, which compares well to the decorrelation time τ_{PC} which is 60 days for the first two POP modes. The values for τ_{PC} compare well to the values for τ_{eig} for lags above about 20 days. Tables 6.1 and 6.3 show a comparison of the characteristic time scales τ_{eig} and τ_{PC}, using different lags and amounts of variance retained.

It was found in this analysis that values of variance kept for the inversion of the covariance matrix above 82% yield robust results. The default variance kept for this analysis was set to 87% along with a lag of 40 days, which is the value used for all figures displayed. The corresponding time scales are highlighted in the tables.

It can be shown that the autocorrelation function looks very similar for the lags and amounts of variance listed in Tables 6.1 and 6.3 (not shown). An uncertainty estimate for the autocorrelation time scales τ_{PC} can be gained from the method proposed in Gerber et al. (2008). An exponential function $e^{-\tau/\tau_{\text{PC}}}$ describing the theoretical decay of the autocorrelation function is fitted to the estimated autocorrelation function

Table 6.2: Same as Table 6.1 but for the pressure weighted data.

	lag 20	**lag 40**	**lag 50**
var 84%	54	62	68
var 87%	54	62	68
var 91%	55	65	70
var 95%	55	64	69

Table 6.3: Comparison of autocorrelation times τ_{PC} computed from the time-series of the POP analysis for varying lag and variance kept for the pseudoinverse. The time scale for the chosen variance and lag is highlighted.

	lag 20	**lag 40**	**lag 50**
var 84%	71	65	66
var 87%	65	60	63
var 91%	71	68	69
var 95%	71	68	69

Table 6.4: Same as Table 6.3 but for the pressure weighted data.

	lag 20	**lag 40**	**lag 50**
var 84%	74	66	67
var 87%	74	66	67
var 91%	74	66	69
var 95%	75	65	68

by minimizing the RMS error between the curves from lag 1 up to the decorrelation time. Following Gerber et al. (2008), the uncertainty in τ_{PC} is estimated based on the standard deviation of the autocorrelation time. This uncertainty was found to be 11 days for the first two POP modes. Applying this uncertainty measure to the characteristic time scales in Tables 6.1 and 6.3 shows that for the presented lags and variances, the time scales lie within the error estimates, both comparing the respective tables as well as within the tables.

It is in addition found that the patterns are robust to changes in spatial resolution. Using only every second data point in pressure and latitude yields the same patterns with robust characteristic times. The patterns and their characteristic time scales also show to be robust to halving the resolution through averaging the data from every two adjacent pressure and latitude levels. In addition, averaging has the positive effect of decreasing the non-diagonal elements of the error covariance matrix $\mathbf{Q}$.

6.5.3 Comparison to Empirical Orthogonal Functions

The presented POP eigenvectors are now compared to the results from EOF analysis. A major difference between EOF and POP analysis is that EOF patterns are sorted according to the variance explained, while POP eigenvectors are sorted according to their persistence, represented by τ_{eig}. The sorting according to variance makes EOF modes sensitive to weighting and normalization, since a different type of variability may be dominant for a weighted pattern. Therefore, the first POP and the first EOF mode do not necessarily show the same pattern.

The EOF modes for the control run are shown in Figure 6-6. The first EOF mode shows a pattern which is similar to the first POP $\mathbf{V}$ matrix. The stratospheric dipole turns out to be the dominant pattern of the modeled zonal wind data in terms of both its persistence and its variance explained. In this case, the first EOF mode explains 52% of the variance while the second mode explains 15% of the variance. The second EOF mode still exhibits a strong stratospheric pattern along with the tropospheric jet wobbling, and looks similar to the imaginary part of the POP matrix with a different surface connection than the first EOF mode, indicating that a similar

type of variability may be explained by the unweighted EOF and POP modes. The autocorrelation times τ_{PC} for the EOF modes are 83 days for the first mode and 49 days for the second mode.

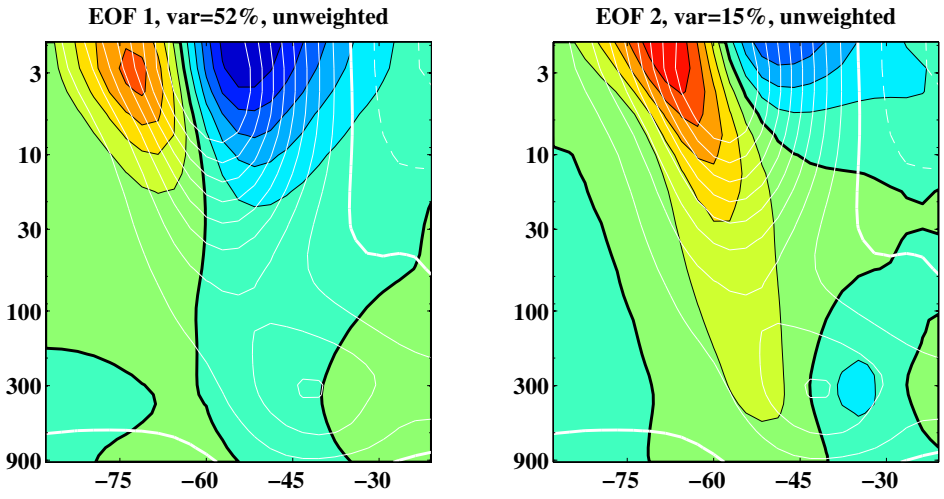

Figure 6-6: Filled contours: EOF contours for mode 1 (left) and mode 2 (right). White contours: Zonal mean zonal wind averaged over the entire run, with a contour interval of $10\,\mathrm{ms}^{-1}$. The zero-wind line is bolded, negative contours are dashed.

6.5.4 Sensitivity to Weighting

Weighting is commonly applied to the computation of EOF modes (Thompson and Wallace, 2000). The latitude weighting adjusts for the decrease in area going poleward and is given by

$$Q(\varphi) = \sqrt{\cos\left(\frac{\pi}{180^\circ}\,\varphi\right)}\,,$$

where φ is latitude [°]. The pressure weighting adjusts for the decrease in mass with height and is given by

$$P(p_j) = \sqrt{\frac{|p_{j+1} - p_j|}{p_0}}\,,$$

where p_j is the pressure at grid point j and p_0 is the pressure at the lowest level. As described in Section 6.3, all modes shown so far are weighted in latitude, however not in pressure. While the latitude weighting by itself marginally changes the EOF modes

(not shown), the pressure weighting strongly confines the patterns to the troposphere, and the weighted modes differ from the the modes without weighting in terms of both amplitude and variability described.

Figure 6-7 shows the EOF modes with applied weighting in both the vertical and latitudinal direction. Comparing Figures 6-6 and 6-7 indicates that the modes are not robust to pressure weighting. While the unweighted first EOF mode dominantly describes a latitudinal wobbling along with a weakening around the stratospheric vortex, the weighted pattern describes a pure weakening. While the pattern is strongly dominated by tropospheric variability, the connection to the troposphere is consistent. While the weighting of the second mode yields a roughly consistent pattern in the stratosphere, it describes a weakening at the location of the tropospheric jet instead of the latitudinal wobbling described by the unweighted pattern, and thereby again a different connection between the stratosphere and the troposphere. The variance explained by the weighted EOF patterns is 48% for the first mode and 20% for the second mode. The decorrelation times are 46 and 8 days for the first and second mode, respectively, which is considerably different from the unweighted patterns (83 and 49 days, respectively).

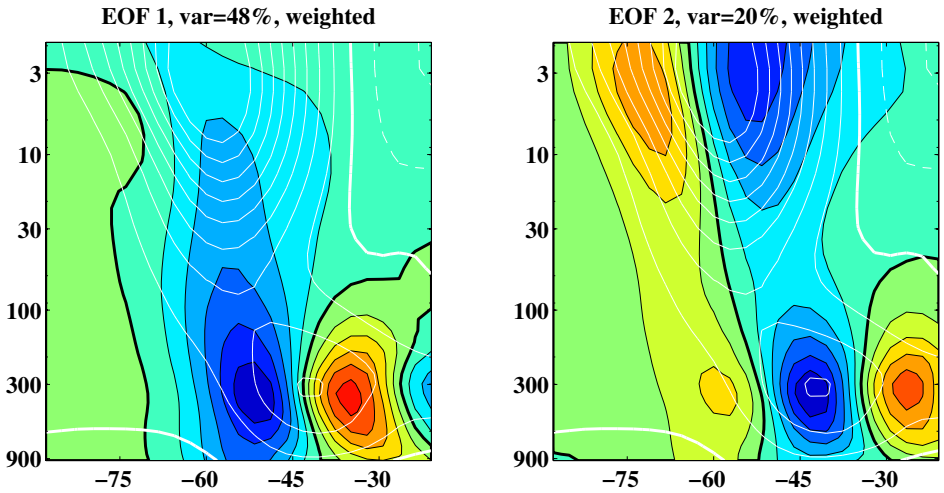

Figure 6-7: Filled contours: Pressure and latitude weighted EOF mode 1 (left) and mode 2 (right). White contours: Zonal mean zonal wind averaged over the entire run, with a contour interval of 10 ms^{-1}. The zero-wind line is bolded, negative contours are dashed.

For comparison, Figures 6-8 and 6-9 show the POP modes for applied pressure and latitude weighting. Comparing to Figures 6-2 and 6-3 shows that the weighting does not affect the patterns as strongly as the EOF patterns. The main characteristics of the patterns are conserved along with the characteristic time scales τ_{eig} and τ_{PC}: Tables 6.2 and 6.4 show the characteristic time scales for the weighted POP modes with varying lag and variance retained. Comparing to the time scales of the unweighted patterns in Tables 6.1 and 6.3 indicates that the effect of the weighting on the characteristic time scale is within the computed errorbars. Similar to EOF patterns, the POP eigenvectors are nearly insensitive to latitude weighting by itself (not shown).

This comparison shows that for including the stratosphere into the analysis, pressure weighting has a major effect on the EOF modes, while POP analysis is more robust to weighting and therefore may be able to overcome the shortcomings of EOF analysis in terms of the bias introduced by pressure weighting.

6.6 Discussion

Principal Oscillation Pattern (POP) analysis has been used in this study to extract the dominant modes of zonal wind variability from a simplified GCM model run exhibiting annular mode - like variability and stratosphere - troposphere coupling. POP modes are designed to represent the time evolution of the dominant structure of a state vector variable. It was found that POP analysis gives a good representation of the dominant variability of the extratropical atmosphere, depicting the stratospheric vortex variability as well as the latitudinal shifts of the tropospheric jet.

While the weighted POP patterns exhibit the expected similarities to the EOF patterns, it was shown that POP analysis has several benefits over EOF analysis, which is commonly used to represent annular mode variability:

- EOF patterns are shown to be strongly dependent on the applied pressure weighting, while POP modes tend to be more independent of applying the same weighting. As opposed to EOF analysis, the POP modes retain their character-

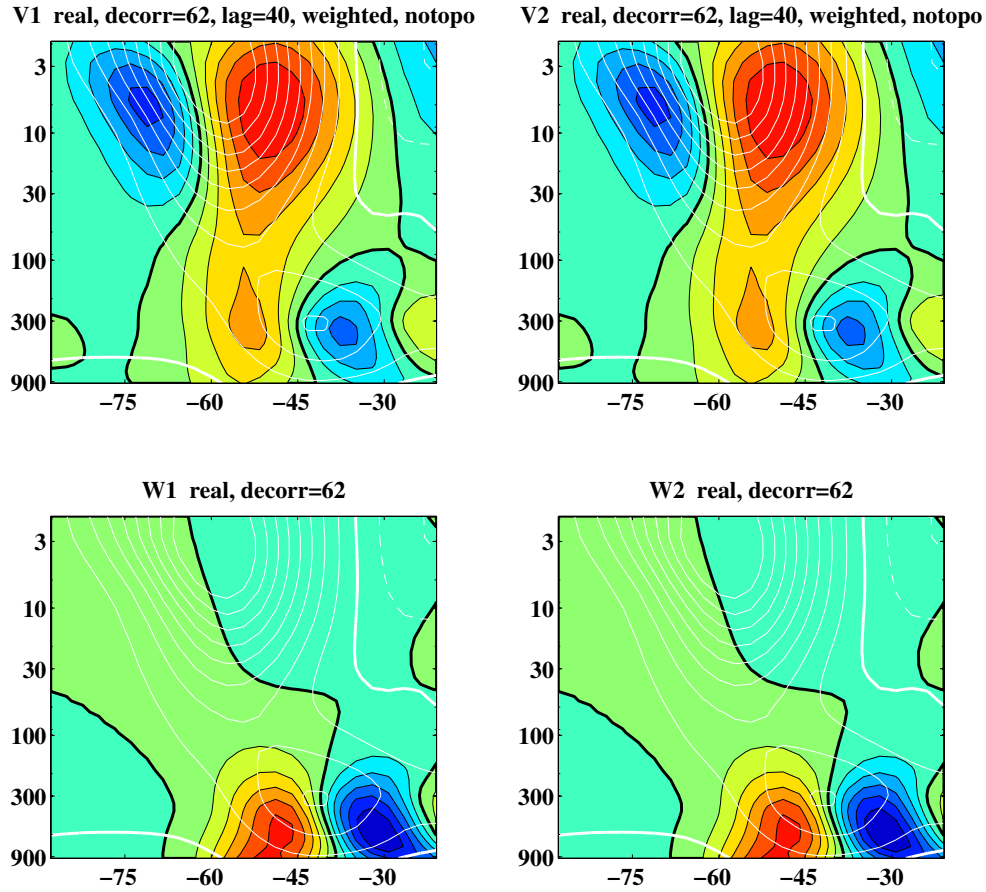

Figure 6-8: Filled contours: Pressure and latitude weighted POP mode 1 (left) and mode 2 (right). White contours: Zonal mean zonal wind averaged over the entire run, with a contour interval of $10\,\mathrm{ms}^{-1}$. The zero-wind line is bolded, negative contours are dashed.

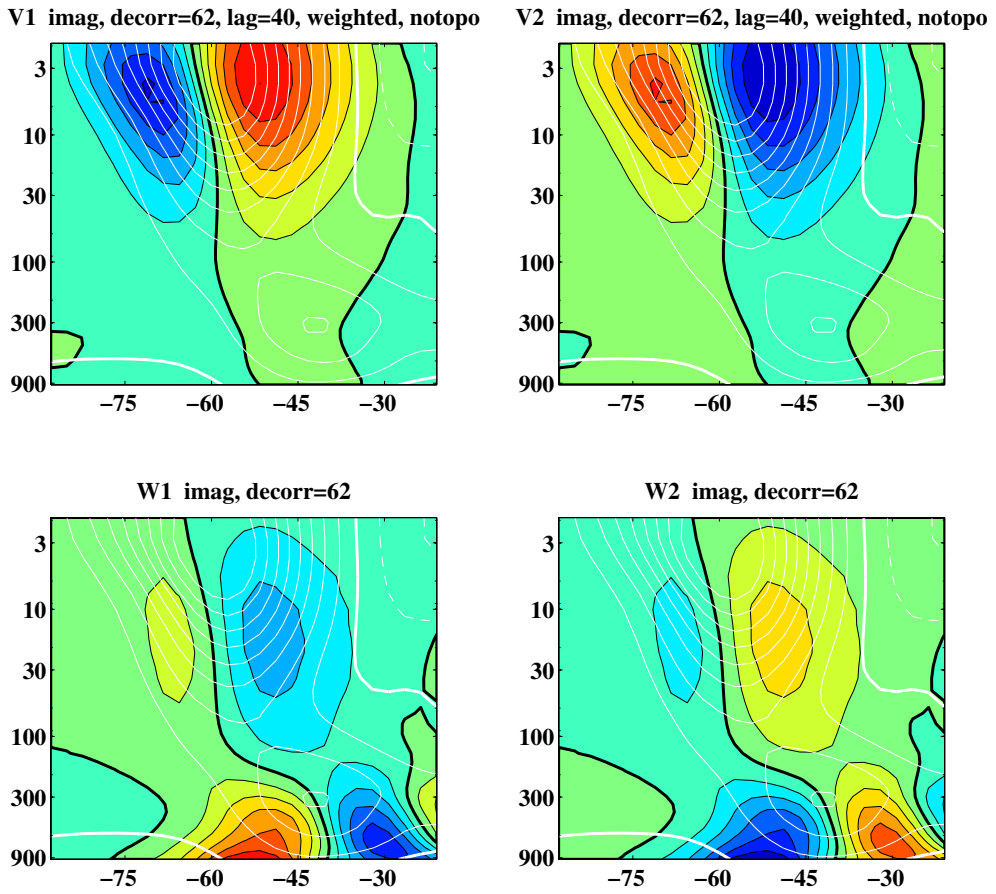

Figure 6-9: Same as Figure 6-8 but for the imaginary part of the pressure weighted POP patterns.

istic time scales and their dominant variability with pressure weighting, making it easier to identify the variability of both the stratosphere and the troposphere and, more importantly, the co-variability between these layers.

- The representations of the annular modes using EOF patterns in the literature are not universal. EOF patterns have to be regressed onto different pressure levels and are used with different weightings, making it difficult to compare the retrieved modes, while no regression has to be used with POP modes.

- While both EOF and POP analysis are able to represent decaying, oscillatory, as well as propagating patterns (EOF analysis by using several patterns, while POP analysis uses only one pattern to represent these structures), POP analysis is in addition able to represent the time dependence of a mode.

The results of this study justify a more extensive use of the POP approach for the analysis of extratropical atmospheric dynamics. POP eigenvectors may help improve the understanding of the interaction between the stratosphere and the troposphere. POP analysis may also help determine the magnitude and influence of a forcing which projects onto the extratropical atmosphere.

However, it was found that the $\mathbf{W}^\top$ modes were not robust in the present analysis, which will limit the ability of POP analysis to predict the response to an imposed forcing. The analysis shown here has not yet included the projection of a forcing onto a pattern in order to quantify and locate the response to a stratospheric forcing in surface climate and the $\mathbf{W}^\top$ pattern are crucial to be able to indicate what pattern a possible forcing will project onto. The analysis of the $\mathbf{W}^\top$ patterns will therefore have to undergo further testing and future work will have to evaluate whether these patterns can be made more robust and if they can be used to project a forcing.

A further extension of the present analysis is the use of reanalysis data. While simple numerical models in combination with statistical analysis provide an essential tool for the understanding of dominant modes, reanalysis data may be a good tool to underline and further explore the potential of POP analysis. This promises to be a valuable extension of the analysis, as simple GCMs may be biased with respect to

overestimating the characteristic time scales (Gerber et al., 2008), which in a Fluctuation - Dissipation framework leads to an overestimation of the response induced by the forcing. In models, the characteristic time scale is sensitive to the inclusion of zonal asymmetries (Gerber and Vallis, 2007), to model resolution in terms of the aspect ratio (Gerber et al., 2008), and to the tropospheric equilibrium profile (Chan and Plumb, 2009).

If further analysis proves the validity and applicability of the described method, it will provide a valuable tool for quantifying the interaction between the extratropical stratosphere and the troposphere in addition to the current use in tropical regions. Purely from a state vector of zonal wind or possibly other related variables such as surface pressure and geopotential, the response to a forcing can be predicted from the eigenvectors of the dynamical operator. This has implications for stratosphere - troposphere interaction as well as for the prediction of surface climate variables.

Being able to define the intrinsic modes of the atmosphere proves to be important for the understanding of the effect of a perturbation in the extratropical atmosphere. A successful definition of the atmospheric modes may also provide information about the mechanism responsible for the interaction in terms of the nature of forcing and response through the respective characteristic time scale. The strong connection between upper atmosphere perturbations and surface climate and the successful representation of these connections by POP analysis encourages further research.

Chapter 7

Conclusions

7.1 Introduction

In this thesis, different aspects of stratospheric sudden warming events and stratosphere - troposphere coupling were investigated. The coupling between the troposphere and the stratosphere is closely connected to the nature and magnitude of stratospheric variability: During times of strong upward wave propagation, the troposphere is able to influence the stratospheric flow in addition to processes which act internally to the stratosphere. A strong and sustained upward wave flux into the stratosphere is able to induce strong stratospheric variability, which can result in major stratospheric sudden warmings. During these events the westerly flow of the stratospheric polar vortex reverses to easterlies, a process which otherwise only radiative processes are able to achieve in the seasonal reversal of the meridional temperature gradient in spring. This indicates that both the tropospheric wave forcing as well as radiative processes are necessary ingredients for the simulation and study of stratospheric variability.

While planetary wave propagation strongly connects the stratosphere to the troposphere before the occurrence of a warming, a tropospheric response to strong and deeply penetrating stratospheric warming events can often be observed in both models and reanalysis data. This tropospheric response is often framed in terms of the annular modes, which describe the dominant internal variability of the extratropical

atmosphere. During times of strong coupling between the stratosphere and the troposphere, the stratospheric mode (describing a weakening and strengthening of the polar vortex) and the tropospheric mode (describing a latitudinal shift of the tropospheric jet) can be shown to be coupled. Stratospheric forcings have been shown to project onto the tropospheric annular modes, indicating that a stratospheric signal may be transformed to strengthen the internal modes of the troposphere.

Comparable to early mechanistic models, the model used in this thesis exhibits the central and necessary ingredients for realistic stratospheric variability: a relaxation to an equilibrium profile on radiative time scales (on the order of several weeks) along with the possibility for the waves to interact with the stratospheric mean flow on time scales much shorter than radiative time scales. Early models generally are 1- or 2-dimensional, often neglecting the meridional degree of freedom. This is satisfying for studying the variability of the zonal mean stratospheric flow in the extratropics. However for observing stratospheric traveling waves or for investigating the impact of stratospheric anomalies onto the troposphere, e.g. in terms of an annular mode response, the possibility of the waves to propagate in three dimensions is essential. The model used here therefore employs the above mentioned critical ingredients for reproducing stratospheric variability along with three spatial dimensions in order to examine stratospheric warmings along with the tropospheric response to a stratospheric forcing. For sudden warmings to occur (i.e. for the simulation of a Northern Hemisphere - like stratosphere), in addition a strong wave source is required at the bottom of the stratosphere or in the model troposphere if the waves are strong enough to reach the stratosphere. The nature and magnitude of the wave forcing was varied for the different questions to be addressed: For the study of traveling waves in the Southern Hemisphere, no stationary waves were forced at the surface and all waves arise purely from internal variability in the model. For studying the Northern Hemisphere, a steady and stationary mountain forcing was employed at the surface. This model setup contains the necessary ingredients for a realistic stratospheric variability.

7.2 Summary of Thesis Work

The overarching question motivating this thesis is the coupling mechanism between the stratosphere and the troposphere. This question was addressed using a simplified general circulation model as described in Chapter 2. As the coupling between the troposphere and stratosphere is particularly strong during Northern Hemisphere winter, the model was run in winter conditions, and tropospheric forcing was varied according to the desired stratospheric variability or tropospheric forcing mechanism. Since the coupling becomes particularly strong during stratospheric sudden warming events, these were simulated using a topographic forcing (Northern Hemisphere - like) in the model. As a first step, the evolution of sudden warming events in the model stratosphere was studied and compared to reanalysis data (Chapter 3). It was found that the model is able to reproduce the essential features of the sudden warming events which are used for the analysis throughout the thesis.

Since upward coupling from the troposphere to the stratosphere is to a major part done by planetary-scale Rossby waves which are forced in the troposphere and propagate into the stratosphere, the model was in a second step used to simulate the tropospheric forcing. While in the Northern Hemisphere, the tropospheric forcing is quite well understood and can be simulated using e.g. topographic (compare to the topography run described in Section 2.2.2) or heat forcings, the generation mechanism and stratospheric effect of traveling planetary waves is less well understood. For this reason, it was studied how traveling waves in the Southern Hemisphere come about and what their impact is on the stratosphere (Chapter 4). The origin of traveling waves was investigated based on the two dominant theories for their origin: interaction of synoptic-scale baroclinic waves aggregating into planetary-scale patterns (Scinocca and Haynes, 1998), as well as direct tropospheric baroclinic instability of planetary-scale waves (Hartmann, 1979). It was found that a typical Southern Hemisphere setup with no surface wave forcing produces traveling waves which are able to propagate into the stratosphere, however the wave flux into the stratosphere is weak and not sustained enough to produce major sudden warmings in the control model

run, consistent with reanalysis data. Only in the absence of synoptic-type variability in the model troposphere is baroclinic instability of planetary-scale waves observed. These instabilities yield a considerable wave flux into the stratosphere leading to Northern Hemisphere - like stratospheric variability including frequent and strong sudden warmings. While it is unclear how applicable this result is to the real atmosphere, it is possible that infrequent periods of weak synoptic wave activity could possibly give rise to planetary-scale disturbances which are able to significantly alter the stratospheric flow.

In a next step, the coupling of the stratosphere and the troposphere was investigated not only for the upward wave coupling, but for the entire evolution of a sudden warming. In order to be able to quantify the respective roles of the wave propagation and the effect of momentum deposition, an angular momentum budget analysis was applied to the atmosphere during the strong coupling effects observed during sudden warmings in order to clarify the exchange mechanisms between the troposphere and the stratosphere. It was found that the upward component of the Eliassen-Palm (EP) flux (proportional to the poleward eddy heat flux) is closely balanced by the Coriolis term at all times in the troposphere, yielding that the effect on the mean flow of the loss of tropospheric wave activity before a sudden warming can only be observed when integrating over a period of several weeks. In the stratosphere, the upward EP flux and the Coriolis term similarly dominate the budget, however they do not balance exactly in magnitude (while they do balance well in time on a daily basis), yielding a significant deceleration of the mean flow, as observed during sudden warmings. However, most of the wave energy provided by the troposphere is converted directly and locally into the induced residual circulation on a daily basis, with the change in mean wind resulting to be a small residual between these two dominant terms. This effect is known as the *Eliassen* response, where a local deposition of momentum is instantaneously transformed into a residual circulation if the duration of the forcing is considerably smaller than the radiative relaxation time, as observed during the sudden warmings simulated in the model run. On the other hand, the upward wave flux and the Coriolis term differ in their remote effect: While the wave flux acts locally

to decelerate the flow and induce the meridional circulation, the residual circulation extends into the tropics, as expected for the time scales of the wave forcing typical for sudden warmings, yielding a strengthening of the mean flow in the subtropical stratosphere as compared to the weakening of the mean flow in the polar regions.

While the tropospheric loss of angular momentum to the stratosphere could be expected to induce a strengthening of the tropospheric mean zonal wind, it is found that this loss translates into a weak negative annular mode response. The annular mode response is reinforced by the stratospheric signal propagating into the troposphere about a month later, i.e. about three weeks after the criterion for a sudden warming is fulfilled in the stratosphere.

It can be concluded that the wave forcing in terms of the EP fluxes cannot be separated from the effect of the residual circulation, as the balance of these terms determines the effect on the mean flow. The small net gain of momentum in the troposphere is translated into a tropospheric annular mode response.

While the momentum budget analysis did not yield the exact mechanism of information exchange between the stratosphere and the troposphere, it gives insight into the balance mechanisms in the stratosphere as well as the tropospheric annular mode response. The annular mode response, however, cannot be readily characterized by the momentum budget and has to be assessed indirectly.

In order to further investigate the effect of the coupling of the annular mode response, a statistical analysis was employed. Tropospheric annular mode variability and thereby the response to a stratospheric forcing is usually framed using Empirical Orthogonal Function (EOF) analysis. However when coupling the annular modes to the stratosphere, the decrease in density has a major impact on the structures detected with EOF modes. Applying a weighting to account for the density difference has a strong impact on the patterns. This thesis introduces Principal Oscillation Pattern (POP) analysis, which is currently used to study remote effects of ocean surface temperatures for seasonal prediction, to stratosphere - troposphere coupling. It is shown that the POP patterns are much less affected by weighting, while they consistently reproduce the annular mode response of the extratropical atmosphere.

7.3 Implications and Proposed Future Work

7.3.1 Implications for Modeling Studies

Truncating a model can lead to important insights into the reasons for why certain processes are either suppressed or dominant in a model atmosphere. The truncated model runs introduced in Chapter 2 and discussed in Chapter 4, originally designed to simplify stratospheric wave propagation as well as the connection to the troposphere, exhibits baroclinic instability of planetary-scale waves after truncating out synoptic waves. This is an important modeling result, i.e. for allowing only a single wave number to propagate (as was e.g. done in early mechanistic models), it will be important not only to eliminate the synoptic waves, but also to eliminate the meridional temperature gradient [see e.g. Scott and Polvani (2004)].

In addition, the finding of baroclinic instability of planetary-scale waves indicates that this process may be important for cases of weak synoptic variability, confirming the results by Hartmann (1979). Concluding, while truncated models are highly artificial, they elucidate processes that may be important in the real atmosphere, but which do not occur on a frequent basis in the real atmosphere or only for specific circumstances.

7.3.2 Implications for Statistical Studies

The successful reproduction of annular mode variability using POP patterns in Chapter 6 encourages a more widespread use of these modes for new applications where weightings and normalizations have to be applied, extending work by Ring and Plumb (2008). The results show that the POP patterns are less sensitive to an applied pressure weighting than the EOF modes, indicating that POP analysis may be suited better to study patterns across areas of large density or other gradients which require a weighting of the data.

7.3.3 Implications for Stratosphere - Troposphere Coupling

The findings in Chapter 5 indicate that the upward wave flux and the induced meridional circulation are tightly coupled and that they can not easily be separated as is often assumed in schematic depictions describing the impact of the upward wave flux on the stratosphere or the impact of sudden warmings on surface climate. The stratosphere and the troposphere are coupled by the very exchange of momentum and they act together to produce both the stratospheric response to an upward wave flux as well as the (albeit weak) tropospheric response to a sudden warming.

However, while the response to a warming is communicated throughout the stratosphere within daily time scales, it does not significantly impact the troposphere until several weeks later. This study confirms earlier studies which indicate that the tropospheric response to an external forcing induces an annular mode response, involving the organizing effect of tropospheric synoptic eddies [confirming e.g. Song and Robinson (2004), Gerber and Polvani (2009), Polvani and Kushner (2002)] and thereby channeling the external forcing into a latitudinal shift of the tropospheric jet, while a direct downward coupling by the induced meridional circulation directly impacting the surface winds (Thompson et al., 2006) is not observed.

For the Southern Hemisphere, Chapter 4 indicates that baroclinic instability of planetary-scale waves is unlikely and therefore not commonly observed in the troposphere. However, the troposphere is baroclinically unstable to planetary-scale disturbances for weak synoptic eddy activity, and the generated planetary waves are able to induce major stratospheric variability. It would be interesting to investigate to what extent baroclinic instability of planetary-scale waves is present during times of weakened synoptic eddy activity. This may be possible by investigating the energy cascade in the troposphere in order to analyze the dominant processes responsible for wave generation in the control run. However it will have to be verified if it is possible to apply the analysis to the short time scales which would be necessary to isolate for planetary-scale baroclinic instability to happen.

It would in addition be interesting to verify to what extent the synoptic eddies need

to be damped in order to allow for more frequent planetary-scale baroclinic instability. From observations, it seems natural to assume that it is very infrequently the case that synoptic waves are weakened enough to allow for planetary-scale instability, however it will have to be verified if this holds for e.g. the Southern Hemisphere warming in September 2002.

While the above findings have clarified the coupling between the stratosphere and the troposphere both during sudden warmings and during weaker stratospheric variability for both hemispheres, the exact mechanism of influence of a lower stratospheric signal on the troposphere is not fully resolved. Even applying the momentum budget derived in this study to every grid point of the model atmosphere (or to any datapoint in an observational dataset in order to investigate the fluxes and their directions) may not conclusively clarify the mechanisms which act to balance the flow, since both passive [Song and Robinson (2004), Charney and Drazin (1961)] and active (Chen and Robinson, 1992) mechanisms are involved in guiding and balancing the observed wave forcing. Tropospheric synoptic variability has a major impact on both stratospheric wave propagation as well as the communication of stratospheric forcing to the troposphere. It can be concluded that the stratosphere and the troposphere are tightly coupled during times of strong stratospheric variability. The observed stratospheric and tropospheric changes before and after a stratospheric sudden warming are small residuals in the constant balance of fluxes between the stratosphere and the troposphere.

Appendix A

Tables

Table A.1: Constants

Constant	Constant name	Value/Units
a	Earth's radius	$6370\,\mathrm{km}$
Ω	rotational constant	$2\pi/24/60/60\ \mathrm{s}^{-1}$
$f(\varphi) = 2\Omega \sin\varphi$	Coriolis parameter	s^{-1}
H	scale height	$7\,\mathrm{km}$
$\rho_0(z)$	density	$\mathrm{kg\,m}^{-3}$
R	gas constant	$287\,\mathrm{J\,K}^{-1}\,\mathrm{kg}^{-1}$
g	gravitational constant	$9.81\ \mathrm{m\,s}^{-2}$

Table A.2: Model Parameters

Parameter	Parameter Name	Unit
h_0	topographic height	m
σ	vertical coordinate	1
ϵ	hemispheric asymmetry	K
γ	stratospheric lapse rate	K km^{-1}
T_{eq}	equilibrium temperature	K

Table A.3: Variables

Variable	Variable Name	Units
φ	latitude	$^\circ$
λ	longitude	$^\circ$
p	pressure	hPa
z	height	m
t	time	s
m	zonal wave number	1
u	zonal wind	ms^{-1}
v	meridional wind	ms^{-1}
w	vertical wind	ms^{-1}
ω	vertical wind in pressure coordinates	hPa s^{-1}
T	temperature	K
θ	potential temperature	K
χ	streamfunction	see text
F_y	meridional EP flux	see text
F_z	vertical EP flux	see text
$\nabla \cdot F$	EP flux divergence	see text
v^*	meridional residual velocity	ms^{-1}
w^*	vertical residual velocity	ms^{-1}
χ^*	residual streamfunction	see text
Φ	geopotential	m^2 s^{-2}
ϕ	geopotential height	m

Bibliography

Alexander, M., L. Matrosova, C. Penland, and J. Scott, 2008: Forecasting pacific SSTs: Linear inverse model predictions of the PDO. *J. Climate*, **21**, 385–402.

Andrews, D., J. Holton, and C. Leony, 1987: *Middle atmosphere dynamics*. Academic Press, San Diego.

Baldwin, M., 2001: Annular modes in global daily surface pressure. *Geophys. Res. Lett.*, **28 (21)**, 4115–4118.

Baldwin, M. and T. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584.

Baldwin, M. and D. Thompson, 2009: A simplified annular mode index based on zonal-mean data. *Q. J. R. Meteorol. Soc.*, **135**, 1661–1672.

Bancalá, S., K. Krüger, and M. Giorgetta, 2012: The preconditioning of major sudden stratospheric warmings. *J. Geophys. Res.*, **117 (D4)**, D04 101.

Butler, A., D. Thompson, and R. Heikes, 2010: The steady-state atmospheric circulation response to climate change-like thermal forcings in a simple general circulation model. *J. Climate*, **23**, 3474–3496.

Chan, C. J. and R. A. Plumb, 2009: The response to stratospheric forcing and its dependence on the state of the troposphere. *J. Atmos. Sci.*, **66 (7)**, 2107–2115.

Charlton, A., A. O'Neill, W. Lahoz, and P. Berrisford, 2005: The splitting of the stratospheric polar vortex in the southern hemisphere, september 2002: Dynamical evolution. *J. Atmos. Sci.: Special Issue on the Southern Hemisphere Sudden Stratospheric Warming of 2002*, **62**, 590–602.

Charlton, A. and L. Polvani, 2007: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *J. Climate*, **20**, 449–469.

Charney, J. and P. Drazin, 1961: Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66 (1)**, 83–109.

Charney, J. and M. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172.

Chen, G. and I. Held, 2007: Phase speed spectra and the recent poleward shift of Southern Hemisphere surface westerlies. *Geophys. Res. Lett.*, **34**, L21 805.

Chen, G. and P. Zurita-Gotor, 2008: The tropospheric jet response to prescribed zonal forcing in an idealized atmospheric model. *J. Atmos. Sci.*, **65 (7)**, 2254.

Chen, P. and W. Robinson, 1992: Propagation of planetary waves between the troposphere and stratosphere. *J. Atmos. Sci.*, **49 (24)**, 2533–2545.

Cooper, F. and P. Haynes, 2008: The response of tropospheric climate to weak perburbations. *Conference Poster*.

Coughlin, K. and L. Gray, 2009: A continuum of sudden stratospheric warmings. *J. Atmos. Sci.*, **66**, 531–540.

Dee, D., et al., 2011: The ERA interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597.

Dickinson, R., 1973: Baroclinic instability of an unbounded zonal shear flow in a compressible atmosphere. *J. Atmos. Sci.*, **30**, 1520–1527.

Dunkerton, T., C.-P. Hsu, and M.E.McIntyre, 1981: Some eulerian and lagrangian diagnostics for a model stratospheric warming. *J. Atmos. Sci.*, **38**, 819–843.

Eliassen, A., 1951: Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophysica Norvegica*, **5 (2)**.

Eliassen, A. and E. Palm, 1961: On the transfer of energy in stationary mountain waves. *Geofys. Publ.*, **22 (3)**, 1–23.

Emanuel, K. A., M. Fantini, and A. J. Thorpe, 1987: Baroclinic instability in an environment of small stability to slantwise moist convection. Part I: Two-dimensional models. *J. Atmos. Sci*, **44 (12)**, 1559–1573.

Esler, J. and N. Matthewman, 2011: Stratospheric sudden warmings as self-tuning resonances. Part II: Vortex displacement events. *J. Atmos. Sci.*, **68**, 2505–2523.

Esler, J., L. Polvani, and R. Scott, 2006: The Antarctic stratospheric sudden warming of 2002: A self-tuned resonance? *Geophy. Res. Lett.*, **33 (L12804)**, 1–5.

Feldstein, S., 2000: The timescale, power spectra, and climate noise properties of teleconnection patterns. *J. Climate*, **13**, 4430–4440.

Garcia, R. R., 1987: On the mean meridional circulation of the middle atmosphere. *J. Atmos. Sci.*, **44 (24)**, 1–11.

Garfinkel, C., T. Shaw, D. Hartmann, and D. Waugh, 2012: Does the Holton-Tan mechanism explain how the Quasi-Biennial Oscillation modulates the Arctic polar vortex? *in press*.

Garfinkel, C. I. and D. L. Hartmann, 2010: Influence of the quasi-biennial oscillation on the North Pacific and El Nino teleconnections. *J. Geophys. Res. - Atmospheres*, **115**, D20 116.

Gerber, E. and G. Vallis, 2007: Eddy–zonal flow interactions and the persistence of the zonal index. *J. Atmos. Sci.*, **64**, 3296 – 3311.

Gerber, E. P., C. Orbe, and L. M. Polvani, 2009: Stratospheric influence on the tropospheric circulation revealed by idealized ensemble forecasts. *Geophys. Res. Lett.*, **36**, L24 801.

Gerber, E. P. and L. M. Polvani, 2009: Stratosphere-troposphere coupling in a relatively simple AGCM: The importance of stratospheric variability. *J. Climate*, **22 (8)**, 1920–1933.

Gerber, E. P., S. Voronin, and L. M. Polvani, 2008: Testing the annular mode autocorrelation time scale in simple atmospheric general circulation models. *Monthly Weather Rev.*, **136 (4)**, 1523.

Gillett, N. and D. Thompson, 2003: Simulation of recent Southern Hemisphere climate change. *Science*, **302**, 273.

Gray, L., S. Phipps, T. Dunkerton, M. Baldwin, E. Drysdale, and M. Allen, 2001: A data study of the influence of the equatorial upper stratosphere on Northern Hemisphere stratospheric sudden warmings. *Q. J. R. Meteorol. Soc.*, **127**, 1985–2003.

Harnik, N. and R. Lindzen, 2001: The effect of reflecting surfaces on the vertical structure and variability of stratospheric planetary waves. *J. Atmos. Sci.*, **58**, 2872–2894.

Hartley, D., J. Villarin, R. Black, and C. Davis, 1998: A new perspective on the dynamical link between the stratosphere and troposphere. *Nature*, **391**, 471–473.

Hartmann, D., 1979: Baroclinic instability of realistic zonal-mean states to planetary waves. *J. Atmos. Sci.*, **36**, 2336–2349.

Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11 012–11 021.

Haynes, P., 2005: Stratospheric dynamics. *Annual Review of Fluid Mechanics*, **37**, 1–31.

Haynes, P., M. McIntyre, T. Shepherd, and C. Marks, 1991: On the 'downward control' of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.*, **48 (4)**, 651–678.

Held, I. and M. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *BAMS*, **75**, 1825–1830.

Hirota, I., T. Hirooka, and M. Shiotani, 1983: Upper stratospheric circulations in the two hemispheres observed by satellites. *Q. J. R. Meteorol. Soc.*, **109 (461)**, 443–454.

Holton, J., P. Haynes, M. McIntyre, and A. Douglass, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.*, **33 (4)**, 403–439.

Holton, J. and H. Tan, 1980: The influence of the equatorial Quasi-Biennial Oscillation on the global circulation at 50mb. *J. Atmos. Sci.*, **37**, 2200–2208.

Holton, J. R. and C. Mass, 1976: Stratospheric vacillation cycles. *J Atmos Sci*, **33**, 2218–2225.

Kalnay, E., et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.*, **77**, 437471.

Kushner, P., I. Held, and T. Delworth, 2001: Southern Hemisphere atmospheric circulation response to global warming. *J. Climate*, **14**, 2238–2249.

Kushner, P. and L. M. Polvani, 2005: A very large, spontaneous stratospheric sudden warming in a simple AGCM: A prototype for the Southern Hemisphere warming of 2002? *J. Atmos. Sci.: Special Issue on the Southern Hemisphere Sudden Stratospheric Warming of 2002*, **62**, 890–897.

Kwok, R. and J. Comiso, 2002: Southern Ocean climate and sea ice anomalies associated with the Southern Oscillation. *J. Climate*, **15**, 487–510.

Labitzke, K., 1977: Interannual variability of the winter stratosphere in the Northern Hemisphere. *Monthly Weather Rev.*, **105**, 762–770.

Labitzke, K., 1979: The major stratospheric warming during January - February 1979. *Beilage zur Berliner Wetterkarte, Meteor. Abh. Berlin*, **23 (B)**.

Labitzke, K., 1981: Stratospheric-mesospheric midwinter disturbances: A summary of observed characteristics. *J. Geophys. Res.*, **86 (C10)**, 9665–9678.

Labitzke, K. and H. van Loon, 1972: The stratosphere in the Southern Hemisphere. *Meteorol. Monogr.*, **13**.

Labitzke, K. and H. van Loon, 1988: Associations between the 11-year solar cycle, the QBO and the atmosphere. Part I: The troposphere and stratosphere in the Northern Hemisphere in winter. *Journal of Atmospheric and Terrestrial Physics*, **50 (3)**, 197–206.

Leith, C., 1975: Climate response and fluctuation dissipation. *J. Atmos. Sci.*, **32**, 2022–2026.

Limpasuvan, V., D. Thompson, and D. Hartmann, 2004: The life cycle of the Northern Hemisphere sudden stratospheric warmings. *J. Climate*, **17**, 2584–2596.

Lorenz, D. and D. Hartmann, 2001: Eddy–zonal flow feedback in the Southern Hemisphere. *J. Atmos. Sci.*, **58**, 3312–3327.

Lorenz, D. and D. Hartmann, 2003: Eddy–zonal flow feedback in the Northern Hemisphere winter. *J. Climate*, **16**, 1212–1227.

Martius, O., L. Polvani, and H. Davies, 2009: Blocking precursors to stratospheric sudden warming events. *Geophys. Res. Lett.*, **36 (L14806)**.

Matsuno, T., 1970: Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *J. Atmos. Sci.*, **27**, 871–883.

Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.*, **28**, 1479–1494.

McInturff, R., 1978: Stratospheric warming: Synoptic, dynamic and general-circulation aspects. *NASA Reference Publication*, **1017**, 1–175.

Mechoso, C. and D. L. Hartmann, 1982: An observational study of traveling planetary waves in the Southern Hemisphere. *J. Atmos. Sci.*, **39**, 1921–1935.

Newman, M. and P. D. Sardeshmukh, 2008: Tropical and stratospheric influences on extratropical short-term climate variability. *J. Climate*, **21 (17)**, 4326.

Newman, P. A. and E. R. Nash, 2005: The unusual Southern Hemisphere stratosphere winter of 2002. *J. Atmos. Sci.: Special Issue on the Southern Hemisphere Sudden Stratospheric Warming of 2002*, **62**, 614–628.

North, G., T. Bell, R. Cahalan, and F. Moeng, 1982: Sampling errors in the estimation of Empirical Orthogonal Functions. *Monthly Weather Rev*, **110**, 699–706.

O'Neill, A., 2003: Stratospheric sudden warmings. *Encyclopedia of Atmospheric Sciences*, 1342–1353.

Orsolini, Y. J., I. T. Kindem, and N. G. Kvamstø, 2011: On the potential impact of the stratosphere upon seasonal dynamical hindcasts of the North Atlantic Oscillation: a pilot study. *Clim. Dyn.*, **36 (3-4)**, 579–588.

Penland, C., 1989: Random forcing and forecasting using principal oscillation pattern analysis. *Monthly Weather Rev*, **117**, 2165–2185.

Penland, C. and P. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2024.

Perlwitz, J. and N. Harnik, 2003: Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection. *J. Climate*, **16**, 3011–3026.

Plumb, R., 1981: Instability of the distorted polar night vortex: A theory of stratospheric warmings. *J. Atmos. Sci.*, **38**, 2514–2531.

Plumb, R., 1982: Zonally symmetric Hough modes and meridional circulations in the middle atmosphere. *J. Atmos. Sci.*, **39**, 983–991.

Plumb, R. A., 1989: On the seasonal cycle of stratospheric planetary waves. *Pure and Applied Geophysics*, **130 (2/3)**, 233–242.

Plumb, R. A., 2010: *Planetary waves and the extratropical winter stratosphere. In: The Stratosphere: Dynamics, Transport and Chemistry*. Polvani, Sobel and Waugh, eds., AGU.

Plumb, R. A. and K. Semeniuk, 2003: Downward migration of extratropical zonal wind anomalies. *J. Geophys. Res.*, **108 (D7)**, 4223.

Polvani, L. and P. Kushner, 2002: Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett*, **29 (7)**, 1114.

Polvani, L. and D. Waugh, 2004: Upward wave activity flux as a precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *J. Climate*, **17**, 3548–3554.

Polvani, L. M., D. W. Waugh, G. Correa, and S.-W. Son, 2011: Stratospheric ozone depletion: the main driver of 20th century atmospheric circulation changes in the Southern Hemisphere. *J. Climate*, **24**, 795–812.

Quiroz, R., 1966: Mid-winter stratospheric warming in the Antarctic revealed by rocket data. *Journal of Applied Meteorology*, **5**, 126–128.

Reichler, T., P. J. Kushner, and L. M. Polvani, 2005: The coupled stratosphere - troposphere response to impulsive forcing from the troposphere. *J. Atmos. Sci.*, **62**, 3337–3352.

Ring, M. and R. A. Plumb, 2007: Forced annular mode patterns in a simple atmospheric general circulation model. *J. Atmos. Sci.*, **64**, 3611.

Ring, M. and R. A. Plumb, 2008: The response of a simplified GCM to axisymmetric forcings: Applicability of the Fluctuation–Dissipation Theorem. *J. Atmos. Sci.*, **65**, 3880.

Ring, M. J., 2008: Annular mode-like responses to external forcings in a simple atmospheric general circulation model. *MIT Doctoral Thesis*.

Scherhag, R., 1952: Die explosionsartige Stratosphärenerwärmung des Spätwinters 1951/52. *Ber. Deut. Wetterdienst*, **38**, 51–63.

Scherhag, R., 1953: Rätselhafte Temperatursprünge in der Stratosphäre. *Die Umschau*, **53**, 646–648.

Schoeberl, M., 1978: Stratospheric warmings - observations and theory. *Reviews of Geophysics and Space Physics*, **16 (4)**, 521–538.

Scinocca, J. F. and P. Haynes, 1998: Dynamical forcing of stratospheric planetary waves by tropospheric baroclinic eddies. *J. Atmos. Sci.*, **55**, 2361–2392.

Scott, R. and L. Polvani, 2004: Stratospheric control of upward wave flux near the tropopause. *Geophys. Res. Lett*, **31**, 1–4.

Scott, R. and L. Polvani, 2006: Internal variability of the winter stratosphere. Part I: Time-independent forcing. *J. Atmos. Sci.*, **63**, 2758–2776.

Shaw, T. and J. Perlwitz, in press: The life cycle of Northern Hemisphere downward wave coupling. *J. Climate*.

Simmons, A., 1974: Baroclinic instability at the winter stratopause. *Q. J. R. Meteorol. Soc.*, **100**, 531—540.

Simmons, A. J. and D. M. Burridge, 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Monthly Weather Rev*, **109**, 758–766.

Son, S.-W., et al., 2008: The impact of stratospheric ozone recovery on the Southern Hemisphere westerly jet. *Science*, **320**, 1486.

Song, Y. and W. Robinson, 2004: Dynamical mechanisms for stratospheric influences on the troposphere. *J. Atmos. Sci.*, **61**, 1711–1725.

Taguchi, M., 2003: Tropospheric response to stratospheric degradation in a simple global circulation model. *J. Atmos. Sci.*, **60**, 1835–1846.

Thompson, D., J. Furtado, and T. G. Shepherd, 2006: On the tropospheric response to anomalous stratospheric wave drag and radiative heating. *J. Atmos. Sci.*, **63**, 2616–2629.

Thompson, D., S. Lee, and M. Baldwin, 2003: Atmospheric processes governing the Northern Hemisphere Annular Mode / North Atlantic Oscillation. *Geophysical Monograph-American Geophysical Union*, **134**, 81–112.

Thompson, D. and J. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25 (9)**, 1297–1300.

Thompson, D. and J. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate*, **13**, 1000–1016.

Trenberth, K. E. and D. P. Stepaniak, 2002: A pathological problem with NCEP reanalyses in the stratosphere. *J. Clim.*, **15**, 690–695.

Uppala, S. M., et al., 2006: The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, **131 (612)**, 2961–3012.

Vallis, G., E. Gerber, P. Kushner, and B. Cash, 2004: A mechanism and simple dynamical model of the North Atlantic Oscillation and annular modes. *J. Atmos. Sci.*, **61**, 264.

von Storch, H., T. Bruns, I. Fischer-Bruns, and K. Hasselmann, 1988: Principal oscillation pattern analysis of the 30-to 60-day oscillation in a general circulation model equatorial troposphere. *J. Geophys. Res*, **93 (D9)**, 11 022–11 036.

Wunsch, C., 2006: *Discrete Inverse and State Estimation Problems with Geophysical Fluid Applications*. Cambridge University Press.

Yoden, S., 1987: Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.*, **44 (13)**, 1723–1733.

Zhang, Y., V. Dymnikov, and J. Wallace, 1997: Sensitivity test of POP system matrices—an application of spectral portrait of a nonsymmetric matrix. *J. Climate*, **10**, 1753–1758.

Zurita-Gotor, P., 2005: Updraft/downdraft constraints for moist baroclinic modes and their implications for the short-wave cutoff and maximum growth rate. *J. Atmos. Sci.*, **62**, 4450–4458.